THE SOLAR SYSTEM

THE SOLAR SYSTEM

THE SOLAR SYSTEM

Volume 1

Archaeoastronomy—Jupiter's Ring System

Editors

David G. Fisher and **Richard R. Erickson**

Lycoming College

Salem Press

Pasadena, California Hackensack, New Jersey

Editor in Chief: Dawn P. Dawson
Editorial Director: Christina J. Moose
Acquisitions Editor: Mark Rehn
Manuscript Editor: Jennifer L. Campbell
Photo Editor: Cynthia Breslin Beres
Production Editor: Joyce I. Buchea
Page Design: James Hutson
Layout: William Zimmerman
Editorial Assistant: Dana Garey

Cover photo: (©Andrea Danti/Dreamstime.com)

Library of Congress Cataloging-in-Publication Data

The solar system / editors, David G. Fisher, Richard R. Erickson.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-58765-530-2 (set : alk. paper) — ISBN 978-1-58765-531-9 (v. 1 : alk. paper) — ISBN 978-1-58765-532-6 (v. 2 : alk. paper) — ISBN 978-1-58765-533-3 (v. 3 : alk. paper) —
1. Solar system. 2. Astronomy. I. Fisher, David G. II. Erickson, Richard R., 1945-
QB501.S627 2010
523.2—dc22

2009013008

First Printing

PRINTED IN THE UNITED STATES OF AMERICA

Publisher's Note

The Solar System covers 180 features of Earth's solar system, including every major body and phenomenon: from the astrophysics of the Sun and the major features of every planet, their satellites, and small bodies such as comets and asteroids, to scientific methodologies, interplanetary phenomena, and related topics in stellar astronomy and cosmology. Authoritative and essential to the core curriculum in astronomy and Earth sciences, these three volumes offer both students and general library patrons detailed basic information on all major aspects of Earth's solar system. No other general reference dedicated to the solar system is this complete and up to date, incorporating the latest perspectives offered by space telescopes, interplanetary probes, and planetary missions.

Scope and Coverage

The essays are designed to meet the needs of both general readers and students enrolled in courses in the Earth sciences (with 25 topics on Earth's geology, geophysics, and astrophysics alone), as well as astronomy, planetology, and cosmology. An older edition, under Salem's Magill's Choice imprint (1998), now more than a decade old, here receives a complete overhaul. For this new edition, we have added 58 topics and have thoroughly expanded every essay, from the text through the bibliographies, to bring the coverage up to date.

The result is the most thorough reference available on our expanding understanding of Earth's universal neighborhood. All of the original 122 essays, as well as the 58 new topics, were reviewed and updated by Professors David G. Fisher and Richard R. Erickson of Lycoming College's Department of Astronomy. They not only have added key developments of the past decade but also have scrutinized and heavily edited the existing text to ensure current accuracy—often increasing the lengths of old essays by more than a third. Such revision and expansion were essential in order to take into account the many planetary and interplanetary missions that have exponentially increased our knowledge of the solar system—from Pioneer and Voyager to the data from Galileo, the Hubble Space Telescope, Cassini-Huygens, NASA's New Horizons program, and the latest Mars rovers.

Organization and Format

Coverage is presented in an A-Z format—from "Archaeoastronomy" and "Asteroids" through "Venus's Volcanoes" to "X-Ray and Gamma-Ray Astronomy"—and is supplemented by hundreds of diagrams and photos. For those wishing instant access to essays grouped by planetary system, a Category List appears in the front of every volume.

The set's essays fall into one or more of the following categories: the Cosmological Context (13), Earth (25), the Jovian System (11), Life in the Solar System (5), Mars (11), Mercury (1), Natural Planetary Satellites (23), the Neptunian System (7), Planets and Planetology (62), the Saturnian System (8), Scientific Methods (13), Small Bodies (17), the Solar System as a Whole (7), the Stellar Context (13), the Sun (20), the Uranian System (7), and Venus (5).

Every essay ranges in length from 2,000 to 5,000 words (3 to 7 pages) and offers not only a complete overview of the topic but also an assessment of knowledge gained, methods of study, or applications. Each essay displays standard ready-reference top matter and subsections:

- **Essay Title:** Topic name, from "Archaeoastronomy" and "Asteroids" through "Mars: Possible Life" to "X-Ray and Gamma-Ray Astronomy."
- **Categories:** Lists scientific subdisciplines relevant to standard undergraduate curricula.
- **Significance:** Summarizes the importance of the topic and current state of our knowledge.
- **Overview:** Rehearses the main facts about the topic.
- **Knowledge Gained *or* Methods of Study *or* Applications:** Headed as appropriate, this section details how the topic is investi-

gated, what scientific knowledge we have accumulated, or the uses of the knowledge we have gained.
- **Context:** Addresses the topic from the larger perspective of the history of solar system science and its relevance for humankind.
- **Byline:** All essays are signed by the scholars who wrote them; these experts are also listed, with their academic affiliations, in the front matter to Volume 1.
- **Further Reading:** An annotated selection of the most important print resources for further study.
- **See also:** Lists cross-references to other essays in *The Solar System* on related topics.

Special Features

Several special resources and finding aids enhance coverage and access to the set's contents. Each volume's front matter includes both an Alphabetical List of Contents (A-Z) and a Category List of Contents (with essays arranged by area of the solar system studied). The set is lavishly illustrated with photos, sidebars, diagrams, and tables. A Glossary, a General Bibliography, a list of Web Sites, and full Subject Index round out the set.

Miscellaneous Notes

The editors have endeavored to follow standard stylistics and usage throughout, using NASA sources in most cases. The words "natural satellite," "satellite" (where unambiguous with technology), and "moon" are used interchangeably. It is important to note that several essays refer to observances of the Sun in optical (visible) light, and we remind readers that one should *never* look directly at the Sun with the naked eye, through a camera, or through sunglasses or other such devices, which offer no protection and can cause significant damage to the eyes.

The Editors and Contributors

Salem Press extends its appreciation to all involved in the development of this work. Special thanks go to the set's editors, David G. Fisher and Richard R. Erickson of Lycoming College's Department of Astronomy, along with their assistant, graduate student Jennifer L. Campbell, who performed much of the manuscript editing. Professor Fisher is Professor of Physics and Astronomy as well as an accomplished spaceflight historian. In the latter capacity, Dr. Fisher served as co-editor of Salem Press's three-volume *USA in Space, Third Edition* (2006). He has published extensively on spaceflight history (concerning not only American robotic and crewed space programs but also Soviet/Russian and other international space projects), physics and physics education, and popular science topics. Professor Erickson is Associate Professor of Astronomy and Physics and Director of the Detwiler Planetarium at Lycoming College. Dr. Erickson has professional interests in and has published articles and papers concerning stellar dynamics, nearby stars, variable stars, galactic structure, relativity and cosmology, impacts and extinctions, and the modeling of the physics of impact events.

The essays were written and are signed by more than eighty scientists, scholars, and other experts, a list of whom will be found in the following pages, accompanied by their academic affiliations. Without their contributions, a project of this magnitude would not be possible.

Contributors

Stephen R. Addison
University of Central Arkansas

Arthur L. Alt
College of Great Falls

Michael S. Ameigh
St. Bonaventure University

Victor R. Baker
University of Arizona

Iona C. Baldridge
Lubbock Christian University

Thomas W. Becker
Webster University

Reta Beebe
New Mexico State University

Timothy C. Beers
Michigan State University

Raymond D. Benge, Jr.
Tarrant County College

Alvin K. Benson
Utah Valley University

John L. Berkley
State University of New York College at Fredonia

Larry M. Browning
South Dakota State University

Michael L. Broyles
Collin County Community College

David S. Brumbaugh
Northern Arizona University

Jessica Lynn Bugno
Lycoming College

Jennifer Campbell
Lycoming College

Dennis Chamberland
Science Writer

D. K. Chowdhury
Indiana University and Purdue University at Fort Wayne

John H. Corbet
Memphis State University

John A. Cramer
Oglethorpe University

Robert L. Cullers
Kansas State University

Joseph Di Rienzi
College of Notre Dame of Maryland

Bruce D. Dod
Mercer University

Dave Dooling
D2 Associates

Steven I. Dutch
University of Wisconsin—Green Bay

John J. Dykla
Loyola University of Chicago

Richard R. Erickson
Lycoming College

Dale C. Ferguson
Baldwin-Wallace College

David G. Fisher
Lycoming College

Richard R. Fisher
National Center for Atmospheric Research

Gerald J. Fishman
National Aeronautics and Space Administration

Michael P. Fitzgerald
Vestal, New York

Dennis R. Flentge
Cedarville College

George J. Flynn
State University of New York College at Plattsburgh

John W. Foster
Illinois State University

Donald R. Franceschetti
Memphis State University

Roberto Garza
San Antonio College

Karl Giberson
Eastern Nazarene College

David Godfrey
National Optical Astronomy Observatory

Gregory A. Good
West Virginia University

C. Alton Hassell
Baylor University

Robert M. Hawthorne, Jr.
Marlboro, Vermont

Paul A. Heckert
Western Carolina University

David Wason Hollar, Jr.
Rockingham Community College

Earl G. Hoover
Science Writer

Hugh S. Hudson
University of California, San Diego

Brian Jones
Science Writer

Richard C. Jones
Texas A&M University

Pamela R. Justice
Collin County Community College

Karen N. Kähler
Pasadena, California

Christopher Keating
University of South Dakota

John P. Kenny
Bradley University

Firman D. King
University of South Florida

Richard S. Knapp
Belhaven College

Narayanan M. Komerath
Georgia Institute of Technology

Kristine Larsen
Central Connecticut State University

Joel S. Levine
National Aeronautics and Space Administration

James C. LoPresto
Edinboro University of Pennsylvania

George E. McCluskey, Jr.
Lehigh University

Michael L. McKinney
University of Tennessee, Knoxville

V. L. Madhyastha
Fairleigh Dickinson University

David W. Maguire
C. S. Mott Community College

Randall L. Milstein
Oregon State University

Brendan Mullan
Pennsylvania State University

Theresa A. Nagy
Pennsylvania State University

Anthony J. Nicastro
West Chester University of Pennsylvania

Divonna Ogier
Oregon Museum of Science and Industry

Steven C. Okulewicz
City University of New York, Hunter College

Satya Pal
New York Institute of Technology

Robert J. Paradowski
Rochester Institute of Technology

Jennifer L. Piatek
Central Connecticut State University

George R. Plitnik
Frostburg State University

Howard L. Poss
Temple University

Gregory J. Retallack
University of Oregon

Clark G. Reynolds
College of Charleston

Mike D. Reynolds
University of North Florida

J. A. Rial
University of North Carolina at Chapel Hill

Charles W. Rogers
Southwestern Oklahoma State University

David M. Schlom
California State University, Chico

Stephen J. Shulik
Clarion University of Pennsylvania

R. Baird Shuman
University of Illinois at Urbana-Champaign

Paul P. Sipiera
William Rainey Harper College

Billy R. Smith, Jr.
Anne Arundel Community College

Roger Smith
Willamette University

Joseph L. Spradley
Wheaton College, Illinois

James L. Whitford-Stark
Sul Ross State University

J. Wayne Wooten
Pensacola Junior College

Clifton K. Yearley
State University of New York at Buffalo

Ivan L. Zabilka
Lexington Kentucky Public Schools

Editors' Introduction

Since ancient times, humans have looked up at the sky and felt both fearsome awe and wondrous amazement. The Sun, Moon, and stars have been worshiped as gods. They have inspired poets, artists, and musicians. Scholars and "ordinary folk" have pondered the nature of what they saw in the sky and our relationship to it.

Probably the first thing to be noticed was that the celestial bodies moved across the sky from east to west, suggesting the Earth was at the center of the universe and everything revolved around us. Most cultures recognized that the stars maintained the same patterns, moving across the sky as groups. Many developed grand stories about these patterns or constellations which became part of their creation mythologies.

In time, it was noted that some of the points of light in the night sky did not move quite the same as most stars did; they "wandered" against the background patterns of the so-called "fixed stars" that remained together in the same constellation. These "wandering stars" or planets (from the Greek word for "wanderer") acquired special significance in many cultures. For example, the Mayans were particularly interested in Venus and kept meticulous records of where it was seen in the sky—sometimes in the east before sunrise and other times in the west after sunset. Some cultures developed rites of worship and built temples devoted to the planets. Some regarded the planets and their movements as portents or divine signs of what would happen here on Earth.

Eclipses and comets also evoked feelings of fear and wonder. Because these visually spectacular events occur only occasionally, those who successfully predicted them could acquire respect and power. There is evidence that one of Stonehenge's functions might have been to predict eclipses; if so, its builders had determined the pattern of eclipses several thousand years ago.

Many cultures tried to explain the nature of the universe (at least what they could see of it) and their place in it. How were the heavens constructed? What kept the clockwork timing of most of the heavens so perfect? What made the planets move so differently? How and why did eclipses occur? Attempts to answer questions like these led to the first models of the universe and theories of cosmology.

A few hundred years B.C.E., Greek philosopher-scientists determined that Earth must be spherical, and they calculated a remarkably accurate value for its circumference, using geometry and a few simple observations. They correctly explained why the Moon went through its phases and how eclipses of the Sun and Moon occurred. They even deduced that the Moon is smaller and the Sun is larger than Earth.

Most Greek models, like those of most other cultures, placed the Earth at the center of all things. Not that some Greeks did not consider the possibility that Earth moved, but such an idea seemed to contradict everything they could perceive. Many centuries would pass and much intense scientific debate would occur before it was accepted that while our existence in the universe may be special, our position is not.

The transformation from an Earth-centered universe to our present picture involved many steps. The first step, led by the likes of Copernicus, Galileo, and Kepler during the 1500's and 1600's, was recognizing that Earth is one of several planets orbiting the Sun. It was not until the 1830's that F. G. W. Struve, F. W. Bessel, and Thomas Henderson independently made the first measurements of stellar parallaxes (which is the only direct way to determine the distances of stars) and thereby confirmed what many had suspected, that stars really are things akin to the Sun but much farther away. Ultimately the Sun was found to be a rather average star. The location of the Sun and solar system in the outer part of the Milky Way galaxy's disk was determined by Harlow Shapley in 1917. A few years later, Edwin Hubble demonstrated that the so-called "spiral nebulae" really are galaxies outside the Milky Way. He then discovered that most galaxy spectra are redshifted, which led to the idea of an expanding universe. At the end of the twentieth century, the expansion of the universe was found to be accelerating, instead of slowing down as was expected. Today, there is serious speculation by some scientists that our universe might be just one of a multitude in a multidimensional "multi-verse."

Observations of planetary motions led to the development of heliocentric models and Sir Isaac Newton's theory of gravity. Early advances in telescope design and construction allowed intriguing but unclear details to be seen on the planets. However, by the middle of the twentieth century, solar system studies had become a backwater of astronomy. Few astronomers then were actively engaged in any sort of planetary or minor body research. (A notable exception was the Dutch-American astronomer Gerard Kuiper.) When one of the editors (Erickson) was an undergraduate at the University of Minnesota in the early 1960's, the textbook used in his first astronomy course had only two chapters consisting of 85 pages (out of a total of 536) that dealt with the planets, their satellites, the asteroids, comets, and meteoroids. (In contrast, current introductory astronomy textbooks typically devote about one-third of their pages to these topics.)

When the other editor (Fisher) was just five years old, he received two cherished books as Christmas presents. Both were from Random House, two selections from the "All About" series. *All About Rockets* and *All About the Universe* were read over and over again. These books describe the status of solar system understanding at the start of the 1960's. The books told about a coming age when humans would send spacecraft out into the solar system to take photographs and make direct measurements of the known planets from Mercury to Pluto, and also examine their moons. The books foretold the adventure of trying to unlock the secrets of the planets and of the great missions to come when humans would extend their reach from the Earth to the planets, maybe even sending people to make discoveries out in space themselves.

The book *All About the Universe* painted a picture of beauty and serenity in both the solar system and the universe around it. Pretty pictures showed the rings of Saturn and the craters of the Moon. Venus was perhaps a planet warmer and wetter than Earth, enjoying a lush, swampy ecosystem filled with plant life. Mars was probably cold and dry, but some form of primitive microbes or plant life might have adapted to live in the harsh environment there. Earth's moon was covered with craters probably created by volcanic action. The moons of the other planets were probably uninteresting balls of ice (with the anticipated exception of Titan about Saturn). Pluto was a small planet, perhaps with some terrestrial characteristics, although the planet would be extremely cold. A very peaceful and safe solar system was described, in which nothing very dramatic happened. All of this was about to change, and a more dynamic solar system soon would be revealed.

The dawn of the space age sparked renewed interest in the solar system. Early Luna (Soviet) and Pioneer (American) flyby missions attempted to reach the Moon in 1959. Only two years after Fisher received *All About the Universe* and *All About Rockets*, a spacecraft named Mariner 2 flew relatively close by Venus and determined that Venus was extraordinarily hot. No lush swamps existed, the surface temperature was hot enough to melt lead, and the atmosphere was thick with carbon dioxide. Then, in 1965, our ideas about Mars changed drastically when Mariner 4 flew by the Red Planet and produced a small set of grainy pictures revealing a dry, cratered surface much more like the Moon than a planet where life might exist.

Rapidly the space program grew in capability and complexity. Now (as of 2009), spacecraft have visited and studied at close range all of the planets in our solar system from Mercury to Neptune, many of their satellites, and several asteroids and comets. More information has been gained about the solar system in a few decades by robotic explorers and space-based observatories than had been learned from millennia of observations from Earth's surface, first with the unaided eye and later with telescopes. Much of what had been believed about solar system bodies and the processes of physical evolution throughout the solar system turned out to be wrong. Craters on the Moon were created largely by impact processes. Earth was (and still is) subject to collisions with comets and asteroids. These large impacts have played a major role in the evolution of the Earth, the development of its atmosphere and oceans, and even the course of biological evolution.

Telescopes were responsible for the first major increase in the storehouse of knowledge and level of understanding about the solar system. In the early 1600's Galileo, using his simple telescopes, found that small bodies orbit around Jupiter (the four Galilean satellites or moons), the Moon has mountains and craters, and the Sun has black spots (sunspots) on it. In 1781, Sir William Herschel discovered Uranus, the first planet not known to the ancients. These early telescopic observations provided the impetus for further investigations of the solar system.

Spacecraft provided the next great leap forward, vastly extending the reach of humanity in investigating our nearest neighborhood, the solar system. However, this is not to say that ground-based telescopes have become obsolete; they continue to provide important clues and new discoveries as well. Solar system studies presently require complementary use of ground-based telescopes, observatories in low-Earth orbit, and spacecraft dedicated to fly out to individual bodies even to the farthest reaches of the solar system.

When the articles for the first edition of *The Solar System* were written, several decades of space exploration had revealed much about our solar system. However, since then there has been such an explosion of information and change in our understanding of the solar system that 58 new articles were written for this second edition. The previous articles have been extensively updated and revised. It has been the editors' job to correct statements in articles that are no longer thought to be true and to update what has been learned since the first edition. For example, when the first edition was published, most scientists believed in the possibility of planets in solar systems other than our own, but none had been definitively detected. At the time of the second edition, approximately three hundred exoplanets have been discovered, and the list continues to grow with each passing year. Furthermore, most scientists have thought that it would be just a matter of technology and time before clear images of planets orbiting other stars could be obtained; now a Jupiter-sized planet orbiting the star Formalhout has been photographed by the Hubble Space Telescope, although it appears only as a tiny dot.

The influx of new information about our solar system and others continues at a furious pace. Only two weeks before this introduction was composed, the French spacecraft COROT detected the closest thing to an Earth-sized planet yet. Named COROT-Exo-7b, this planet's existence was revealed by observing a dip in the planet's star's brightness as the planet transited in front of the star. The data suggest that this exoplanet is between two and four times the size of our Earth. This planet's "year" is less than one Earth day (approximately 20 hours), which means it orbits extraordinarily close to its star; consequently, this planet must be extremely hot (possibly 1,800 kelvins) and would share very little in common with Earth. A spacecraft named Kepler was launched on March 6, 2009. This Earth-orbiting observatory will examine a far greater number of distant stars by the transit methodology to search for planets—both smaller and larger than Earth—orbiting them. The time may well be near when an Earth-like planet will be imaged and its atmosphere spectroscopically investigated. Famed astrophysicist Michio Kaku commented in regard to the COROT-Exo-7b discovery that humanity could experience an existential shock when we first find a planet with conditions similar to those on Earth orbiting a star within its habitable zone, thus demonstrating that there is a place besides Earth where life might have developed.

Closer to home, as a result of discoveries of bodies similar to and farther out in the solar system than Pluto, the definition of "planet" itself was altered at a meeting of the International Astronomical Union in 2006; Pluto was "demoted" to the status of dwarf planet, becoming instead the first member in a new category of objects called plutoids, leaving our solar system (by convention) now with eight recognized planets. The worldview of solar system objects continues to undergo major changes with increasingly ambitious space missions. The Mars

Phoenix Lander provided the first direct evidence of subsurface water in the Red Planet's northern polar region. For several decades the search for possible primitive life on Mars has been based on the paradigm of "follow the water." Shortly after the Mars Phoenix confirmation that subsurface permafrost exists on Mars, a new discovery of methane in the Martian atmosphere held the promise of a different method for looking for life: Methane replenishment at Mars might indicate a biological origin and may shed light on the methane abundance in the atmosphere on Saturn's moon Titan.

At this writing, the issue of global climate change on Earth remains controversial. To truly understand what might be happening on Earth and the implications for life on Earth, changes in our atmosphere and climate need to be addressed in the broader context of comparative planetology. Venus, Earth, and Mars appear to have had very similar conditions in the past. Why did Venus lose most of its water, develop a thick carbon dioxide atmosphere, and heat up to the point where the surface temperature is enough to melt lead? Why did Earth retain its water, and how did the water lead to conditions where life as we know it could develop? How did Mars lose the vast majority of its surface water and develop a dry, arid surface with only a very thin atmosphere of carbon dioxide? Answers to these questions may lead to a better determination of the direction in which climate change on Earth might be headed.

Global climate change is but one possible threat to Earth. Whereas in the past the solar system had been considered to be a serene, quiet place, it is now understood that violent impacts have played a major role in the evolution of planetary surfaces and even on the evolution of life here on Earth. Some large impacts may have helped spread and promote the development of life, while others caused mass extinctions. An asteroid or comet impact remains a real threat today. Insufficient research is being done to detect those objects which might be heading "our" way. Fifty thousand years ago a small body about half as large as a football field created the impressive Meteor Crater near Flagstaff, Arizona. Compared to the asteroid that probably killed off the dinosaurs 65 million years ago, the one that produced Meteor Crater was a baby. While investigation of potential global climate change is important, a greater threat to life on Earth almost certainly is impact of a large body from space. With the promise of technology, such a natural disaster could be avoided as a result of direct intervention, but only if the threat from a specific body is identified well in advance of its impact.

The field of planetary research will surely remain a rich one, and questions raised by today's investigations may be answered in the relatively near future. Also in the years to come, new questions will be asked in response to information being obtained currently. A spacecraft, New Horizons, is on its way toward Pluto, and it will return images of that body and other Kuiper Belt objects. Efforts are now under way to expand the human presence in space beyond low-Earth orbit, to which it has been restricted since the conclusion of the Apollo program in 1972. The technology for a return to the Moon to stay and to eventually send human beings to Mars is being developed. If these plans materialize, humans may well establish bases upon the lunar surface from which extensive investigations of lunar geology or astronomical observations can be performed on a regular basis. If humans venture out to Mars, questions that numerous robotic probes have sought to address might finally be answered. Mars Exploration Rover principal investigator Dr. Steve Squyres, in his book about the plucky Spirit and Opportunity rovers, remarked that he, as a planetary geologist with trained eyes and hands, could have done all the work of the rovers in just a week's worth of his own surface investigations on Mars. Perhaps soon, scientists will make a better determination of the orbit of 99942 Apophis, an asteroid that will pass very close to (but not hit) the Earth in 2029. It is feasible that, by then, technology will be capable of using gravitational interactions to "tractor" that asteroid away, eliminating its threats to Earth and leading to a planetary defense system against future impacts.

Only the future will tell.

David G. Fisher, Lycoming College
Richard R. Erickson, Lycoming College
February 15, 2009

Contents

Publisher's Note v
Contributors vii
Editor's Introduction ix
Units of Measure. xv
Alphabetical List of Contents xxi
Category List of Contents xxv

Archaeoastronomy 1
Asteroids 6
Auroras. 17

Big Bang 24
Brown Dwarfs 31

Callisto 35
Ceres 40
Comet Halley 44
Comet Shoemaker-Levy 9 52
Comets 56
Coordinate Systems 65
Coronal Holes and Coronal Mass
Ejections. 69
Cosmic Rays 73
Cosmology 78

Dwarf Planets 85

Earth-Moon Relations 90
Earth-Sun Relations. 94
Earth System Science 99
Earth's Age 103
Earth's Atmosphere 108
Earth's Composition 115
Earth's Core 121
Earth's Core-Mantle Boundary 125
Earth's Crust 129
Earth's Crust-Mantle Boundary: The
Mohorovičić Discontinuity 135
Earth's Differentiation 139
Earth's Magnetic Field: Origins 145
Earth's Magnetic Field: Secular
Variation 149
Earth's Magnetic Field at
Present. 154
Earth's Magnetosphere 158
Earth's Mantle 162
Earth's Oceans 166
Earth's Origin 173
Earth's Rotation 177
Earth's Shape. 183
Earth's Structure. 186
Eclipses 190
Electromagnetic Radiation:
Nonthermal Emissions 197
Electromagnetic Radiation:
Thermal Emissions 201
Enceladus. 204
Eris and Dysnomia 208
Europa 213
Extrasolar Planets 220
Extrasolar Planets: Detection
Methods 227
Extraterrestrial Life in the Solar
System 231

Gamma-Ray Bursters 238
Ganymede 244
General Relativity 248
Gravity Measurement 252
Greenhouse Effect 260

Habitable Zones 266
Hertzsprung-Russell Diagram 269

Iapetus 275
Impact Cratering 279
Infrared Astronomy 286
Interplanetary Environment. 295
Interstellar Clouds and the
Interstellar Medium 299
Io 304

Jovian Planets 311
Jupiter's Atmosphere 318
Jupiter's Great Red Spot 324
Jupiter's Interior 331
Jupiter's Magnetic Field and
Radiation Belts 335
Jupiter's Ring System 341

Units of Measure

Common prefixes for metric units—which may apply in more cases than shown above—include *giga-* (1 billion times the unit), *mega-* (one million times), *kilo-* (1,000 times), *hecto-* (100 times), *deka-* (10 times), *deci-* (0.1 times, or one tenth the unit), *centi-* (0.01, or one hundredth), *milli-* (0.001, or one thousandth), and *micro-* (0.0001, or one millionth).

Unit	***Quantity***	***Symbol***	***Equivalents***
Acre	Area	ac	43,560 square feet 4,840 square yards 0.405 hectare
Ampere	Electric current	A *or* amp	1.00016502722949 international ampere 0.1 biot *or* abampere
Angstrom	Length	Å	0.1 nanometer 0.0000001 millimeter 0.000000004 inch
Astronomical unit	Length	AU	92,955,807 miles 149,597,871 kilometers (mean Earth-Sun distance)
Barn	Area	b	10^{-28} meters squared (approx. cross-sectional area of 1 uranium nucleus)
Barrel (dry, for most produce)	Volume/capacity	bbl	7,056 cubic inches; 105 dry quarts; 3.281 bushels, struck measure
Barrel (liquid)	Volume/capacity	bbl	31 to 42 gallons
British thermal unit	Energy	Btu	1055.05585262 joule
Bushel (U.S., heaped)	Volume/capacity	bbl	2,747.715 cubic inches 1.278 bushels, struck measure
Bushel (U.S., struck measure)	Volume/capacity	bsh *or* bu	2,150.42 cubic inches 35.238 liters
Candela	Luminous intensity	cd	1.09 hefner candle
Celsius	Temperature	C	1° centigrade
Centigram	Mass/weight	cg	0.15 grain
Centimeter	Length	cm	0.3937 inch
Centimeter, cubic	Volume/capacity	cm^3	0.061 cubic inch
Centimeter, square	Area	cm^2	0.155 square inch
Coulomb	Electric charge	C	1 ampere second

Unit	***Quantity***	***Symbol***	***Equivalents***
Cup	Volume/capacity	C	250 milliliters 8 fluid ounces 0.5 liquid pint
Deciliter	Volume/capacity	dl	0.21 pint
Decimeter	Length	dm	3.937 inches
Decimeter, cubic	Volume/capacity	dm^3	61.024 cubic inches
Decimeter, square	Area	dm^2	15.5 square inches
Dekaliter	Volume/capacity	dal	2.642 gallons 1.135 pecks
Dekameter	Length	dam	32.808 feet
Dram	Mass/weight	dr *or* dr avdp	0.0625 ounce 27.344 grains 1.772 grams
Electron volt	Energy	eV	$1.5185847232839 \times 10^{-22}$ Btus $1.6021917 \times 10^{-19}$ joules
Fermi	Length	fm	1 femtometer 1.0×10^{-15} meters
Foot	Length	ft *or* ′	12 inches 0.3048 meter 30.48 centimeters
Foot, cubic	Volume/capacity	ft^3	0.028 cubic meter 0.0370 cubic yard 1,728 cubic inches
Foot, square	Area	ft^2	929.030 square centimeters
Gallon (British Imperial)	Volume/capacity	gal	277.42 cubic inches 1.201 U.S. gallons 4.546 liters 160 British fluid ounces
Gallon (U.S.)	Volume/capacity	gal	231 cubic inches 3.785 liters 0.833 British gallon 128 U.S. fluid ounces
Giga-electron volt	Energy	GeV	$1.6021917 \times 10^{-10}$ joule
Gigahertz	Frequency	GHz	—
Gill	Volume/capacity	gi	7.219 cubic inches 4 fluid ounces 0.118 liter
Grain	Mass/weight	gr	0.037 dram 0.002083 ounce 0.0648 gram

Unit	***Quantity***	***Symbol***	***Equivalents***
Gram	Mass/weight	g	15.432 grains 0.035 avoirdupois ounce
Hectare	Area	ha	2.471 acres
Hectoliter	Volume/capacity	hl	26.418 gallons 2.838 bushels
Hertz	Frequency	Hz	$1.08782775707767 \times 10^{-10}$ cesium atom frequency
Hour	Time	h	60 minutes 3,600 seconds
Inch	Length	in *or* ″	2.54 centimeters
Inch, cubic	Volume/capacity	in^3	0.554 fluid ounce 4.433 fluid drams 16.387 cubic centimeters
Inch, square	Area	in^2	6.4516 square centimeters
Joule	Energy	J	$6.2414503832469 \times 10^{18}$ electron volt
Joule per kelvin	Heat capacity	J/K	$7.24311216248908 \times 10^{22}$ Boltzmann constant
Joule per second	Power	J/s	1 watt
Kelvin	Temperature	K	–272.15 Celsius
Kilo-electron volt	Energy	keV	$1.5185847232839 \times 10^{-19}$ joule
Kilogram	Mass/weight	kg	2.205 pounds
Kilogram per cubic meter	Mass/weight density	kg/m^3	$5.78036672001339 \times 10^{-4}$ ounces per cubic inch
Kilohertz	Frequency	kHz	—
Kiloliter	Volume/capacity	kl	—
Kilometer	Length	km	0.621 mile
Kilometer, square	Area	km^2	0.386 square mile 247.105 acres
Light-year (distance traveled by light in one Earth year)	Length/distance	lt-yr	5,878,499,814,275.88 miles 9.46×10^{12} kilometers
Liter	Volume/capacity	L	1.057 liquid quarts 0.908 dry quart 61.024 cubic inches
Mega-electron volt	Energy	MeV	—
Megahertz	Frequency	MHz	—

Unit	***Quantity***	***Symbol***	***Equivalents***
Meter	Length	m	39.37 inches
Meter, cubic	Volume/capacity	m^3	1.308 cubic yards
Meter per second	Velocity	m/s	2.24 miles per hour 3.60 kilometers per hour
Meter per second per second	Acceleration	m/s^2	12,960.00 kilometers per hour per hour 8,052.97 miles per hour per hour
Meter, square	Area	m^2	1.196 square yards 10.764 square feet
Metric. *See* unit name			
Microgram	Mass/weight	mcg *or* μg	0.000001 gram
Microliter	Volume/capacity	μl	0.00027 fluid ounce
Micrometer	Length	μm	0.001 millimeter 0.00003937 inch
Mile (nautical international)	Length	mi	1.852 kilometers 1.151 statute miles 0.999 U.S. nautical miles
Mile (statute or land)	Length	mi	5,280 feet 1.609 kilometers
Mile, square	Area	mi^2	258.999 hectares
Milligram	Mass/weight	mg	0.015 grain
Milliliter	Volume/capacity	ml	0.271 fluid dram 16.231 minims 0.061 cubic inch
Millimeter	Length	mm	0.03937 inch
Millimeter, square	Area	mm^2	0.002 square inch
Minute	Time	m	60 seconds
Mole	Amount of substance	mol	6.02×10^{23} atoms or molecules of a given substance
Nanometer	Length	nm	1,000,000 fermis 10 angstroms 0.001 micrometer 0.00000003937 inch
Newton	Force	N	$6.14124095407198 \times 10^{25}$ atomic weight 0.224808943099711 pound force 0.101971621297793 kilogram force 100,000 dynes

Unit	*Quantity*	*Symbol*	*Equivalents*
Newton meter	Torque	N·m	0.7375621 foot-pound
Ounce (avoirdupois)	Mass/weight	oz	28.350 grams 437.5 grains 0.911 troy or apothecaries' ounce
Ounce (troy)	Mass/weight	oz	31.103 grams 480 grains 1.097 avoirdupois ounces
Ounce (U.S., fluid or liquid)	Mass/weight	oz	1.805 cubic inch 29.574 milliliters 1.041 British fluid ounces
Parsec	Length	pc	30,856,775,876,793 kilometers 19,173,511,615,163 miles
Peck	Volume/capacity	pk	8.810 liters
Pint (dry)	Volume/capacity	pt	33.600 cubic inches 0.551 liter
Pint (liquid)	Volume/capacity	pt	28.875 cubic inches 0.473 liter
Pound (avoirdupois)	Mass/weight	lb	7,000 grains 1.215 troy or apothecaries' pounds 453.59237 grams
Pound (troy)	Mass/weight	lb	5,760 grains 0.823 avoirdupois pound 373.242 grams
Quart (British)	Volume/capacity	qt	69.354 cubic inches 1.032 U.S. dry quarts 1.201 U.S. liquid quarts
Quart (U.S., dry)	Volume/capacity	qt	67.201 cubic inches 1.101 liters 0.969 British quart
Quart (U.S., liquid)	Volume/capacity	qt	57.75 cubic inches 0.946 liter 0.833 British quart
Rod	Length	rd	5.029 meters 5.50 yards
Rod, square	Area	rd^2	25.293 square meters 30.25 square yards 0.00625 acre
Second	Time	s *or* sec	$\frac{1}{60}$ minute $\frac{1}{3600}$ hour

Unit	***Quantity***	***Symbol***	***Equivalents***
Tablespoon	Volume/capacity	T *or* tb	3 teaspoons 4 fluid drams
Teaspoon	Volume/capacity	t *or* tsp	0.33 tablespoon 1.33 fluid drams
Ton (gross or long)	Mass/weight	t	2,240 pounds 1.12 net tons 1.016 metric tons
Ton (metric)	Mass/weight	t	1,000 kilograms 2,204.62 pounds 0.984 gross ton 1.102 net tons
Ton (net or short)	Mass/weight	t	2,000 pounds 0.893 gross ton 0.907 metric ton
Volt	Electric potential	V	1 joule per coulomb
Watt	Power	W	1 joule per second 0.001 kilowatt $2.84345136093995 \times 10^{-4}$ ton of refrigeration
Yard	Length	yd	0.9144 meter
Yard, cubic	Volume/capacity	yd^3	0.765 cubic meter
Yard, square	Area	yd^2	0.836 square meter

Alphabetical List of Contents

Volume 1

Publisher's Note v
Contributors vii
Editor's Introduction ix
Contents xiii
Units of Measure xv
Alphabetical List of Contents xxi
Category List of Contents xxv

Archaeoastronomy 1
Asteroids 6
Auroras 17

Big Bang 24
Brown Dwarfs 31

Callisto 35
Ceres 40
Comet Halley 44
Comet Shoemaker-Levy 9 52
Comets 56
Coordinate Systems 65
Coronal Holes and Coronal Mass Ejections 69
Cosmic Rays 73
Cosmology 78

Dwarf Planets 85

Earth-Moon Relations 90
Earth-Sun Relations 94
Earth System Science 99
Earth's Age 103
Earth's Atmosphere 108
Earth's Composition 115
Earth's Core 121
Earth's Core-Mantle Boundary 125
Earth's Crust 129
Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity 135
Earth's Differentiation 139
Earth's Magnetic Field: Origins 145
Earth's Magnetic Field: Secular Variation 149
Earth's Magnetic Field at Present 154
Earth's Magnetosphere 158
Earth's Mantle 162
Earth's Oceans 166
Earth's Origin 173
Earth's Rotation 177
Earth's Shape 183
Earth's Structure 186
Eclipses 190
Electromagnetic Radiation: Nonthermal Emissions 197
Electromagnetic Radiation: Thermal Emissions 201
Enceladus 204
Eris and Dysnomia 208
Europa 213
Extrasolar Planets 220
Extrasolar Planets: Detection Methods 227
Extraterrestrial Life in the Solar System 231

Gamma-Ray Bursters 238
Ganymede 244
General Relativity 248
Gravity Measurement 252
Greenhouse Effect 260

Habitable Zones 266
Hertzsprung-Russell Diagram 269

Iapetus 275
Impact Cratering 279
Infrared Astronomy 286
Interplanetary Environment 295
Interstellar Clouds and the Interstellar Medium 299
Io 304

Jovian Planets 311
Jupiter's Atmosphere 318
Jupiter's Great Red Spot 324
Jupiter's Interior 331
Jupiter's Magnetic Field and Radiation Belts 335
Jupiter's Ring System 341

Volume 2

Contents xxxv
Units of Measure xxxvii
Alphabetical List of Contents xliii
Category List of Contents xlvii

Jupiter's Satellites 345

Kuiper Belt 352

Life's Origins 355
Lunar Craters 360
Lunar History 368
Lunar Interior 374
Lunar Maria 378
Lunar Regolith Samples 384
Lunar Rocks 389
Lunar Surface Experiments 394

Main Sequence Stars 398
Mars: Illusions and Conspiracies 401
Mars: Possible Life 405
Mars: Surface Experiments 409
Mars's Atmosphere 415
Mars's Craters 421
Mars's Polar Caps 428
Mars's Satellites 437
Mars's Valleys 441
Mars's Volcanoes 446
Mars's Water 452
Mercury 460
Meteorites: Achondrites 468
Meteorites: Carbonaceous Chondrites 472
Meteorites: Chondrites 477
Meteorites: Nickel-Irons 481
Meteorites: Stony Irons 486
Meteoroids from the Moon and Mars 490
Meteors and Meteor Showers 495
Milky Way 501
Miranda 506
Nemesis and Planet X 510
Neptune's Atmosphere 516
Neptune's Great Dark Spots 520
Neptune's Interior 524
Neptune's Magnetic Field 527
Neptune's Ring System 530
Neptune's Satellites 535
Neutrino Astronomy 538
Novae, Bursters, and X-Ray Sources 543
Nuclear Synthesis in Stars 549

Oort Cloud 555
Optical Astronomy 558

Planetary Atmospheres 564
Planetary Classifications 572
Planetary Formation 575
Planetary Interiors 582
Planetary Magnetospheres 589
Planetary Orbits 594
Planetary Orbits: Couplings and Resonances 601
Planetary Ring Systems 605
Planetary Rotation 612
Planetary Satellites 617
Planetary Tectonics 625
Planetology: Comparative 630
Planetology: Venus, Earth, and Mars . . . 637
Pluto and Charon 641
Protostars 649
Pulsars 653

Radio Astronomy 659
Red Dwarf Stars 665
Red Giant Stars 669

Saturn's Atmosphere 674
Saturn's Interior 683
Saturn's Magnetic Field 686
Saturn's Ring System 693

Volume 3

Contents lv
Units of Measure lvii
Alphabetical List of Contents lxiii
Category List of Contents lxvii

Saturn's Satellites 701
Search for Extraterrestrial Intelligence 710
Solar Chromosphere 713
Solar Corona 717
Solar Evolution 723
Solar Flares. 727
Solar Geodesy. 732
Solar Infrared Emissions 736
Solar Interior 739
Solar Magnetic Field 743
Solar Photosphere 748
Solar Radiation 751
Solar Radio Emissions 756
Solar Seismology 759
Solar Structure and Energy 763
Solar System: Element Distribution. . . . 767
Solar System: Origins 773
Solar Ultraviolet Emissions 778
Solar Variability 784
Solar Wind 787
Solar X-Ray Emissions 792
Space-Time: Distortion by Gravity 795
Space-Time: Mathematical Models 801
Stellar Evolution 805
Sunspots 811
Supernovae 816

Telescopes: Ground-Based 824
Telescopes: Space-Based 832
Terrestrial Planets 843
Thermonuclear Reactions in Stars 847
Titan 852
Triton 857

Ultraviolet Astronomy 862
Universe: Evolution 867
Universe: Expansion 874
Universe: Structure 879
Uranus's Atmosphere 886
Uranus's Interior 890
Uranus's Magnetic Field 893
Uranus's Rings 896
Uranus's Satellites 901
Uranus's Tilt 906

Van Allen Radiation Belts 910
Venus's Atmosphere 914
Venus's Craters. 921
Venus's Surface Experiments 925
Venus's Surface Features 929
Venus's Volcanoes 938

White and Black Dwarfs 942

X-Ray and Gamma-Ray Astronomy 947

Appendixes
Glossary. 957
General Bibliography 992
Web Sites 1011

Index
Subject Index 1017

Category List of Contents

LIST OF CATEGORIES

The Cosmological Context xxv
Earth xxv
The Jovian System xxv
Life in the Solar System xxvi
Mars xxvi
Mercury xxvi
Natural Planetary Satellites xxvi
The Neptunian System xxvi
Planets and Planetology xxvi
The Saturnian System xxvii
Scientific Methods xxvii
Small Bodies xxvii
The Solar System as a Whole. xxvii
The Stellar Context xxvii
The Sun xxviii
The Uranian System xxviii
Venus xxviii

Asteroids. *See* **Small Bodies**

Comets. *See* **Small Bodies**

The Cosmological Context
Big Bang, 24
Cosmic Rays, 73
Cosmology, 78
Electromagnetic Radiation: Nonthermal Emissions, 197
Electromagnetic Radiation: Thermal Emissions, 201
General Relativity, 248
Interstellar Clouds and the Interstellar Medium, 299
Milky Way, 501
Space-Time: Distortion by Gravity, 795
Space-Time: Mathematical Models, 801
Universe: Evolution, 867
Universe: Expansion, 874
Universe: Structure, 879

Dwarf planets. *See* **Small Bodies**

Earth
Auroras, 17
Earth-Moon Relations, 90
Earth-Sun Relations, 94
Earth System Science, 99
Earth's Age, 103
Earth's Atmosphere, 108
Earth's Composition, 115
Earth's Core, 121
Earth's Core-Mantle Boundary, 125
Earth's Crust, 129
Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity, 135
Earth's Differentiation, 139
Earth's Magnetic Field at Present, 154
Earth's Magnetic Field: Origins, 145
Earth's Magnetic Field: Secular Variation, 149
Earth's Magnetosphere, 158
Earth's Mantle, 162
Earth's Oceans, 166
Earth's Origin, 173
Earth's Rotation, 177
Earth's Shape, 183
Earth's Structure, 186
Eclipses, 190
Greenhouse Effect, 260
Van Allen Radiation Belts, 910

The Jovian System
Callisto, 35
Europa, 213
Ganymede, 244
Io, 304
Jovian Planets, 311
Jupiter's Atmosphere, 318
Jupiter's Great Red Spot, 324
Jupiter's Interior, 331
Jupiter's Magnetic Field and Radiation Belts, 335
Jupiter's Ring System, 341
Jupiter's Satellites, 345

Jupiter. *See* **The Jovian System**

Life in the Solar System
Extraterrestrial Life in the Solar System, 231
Habitable Zones, 266
Life's Origins, 355
Mars: Possible Life, 405
Search for Extraterrestrial Intelligence, 710

Mars
Mars: Illusions and Conspiracies, 401
Mars: Possible Life, 405
Mars: Surface Experiments, 409
Mars's Atmosphere, 415
Mars's Craters, 421
Mars's Polar Caps, 428
Mars's Satellites, 437
Mars's Valleys, 441
Mars's Volcanoes, 446
Mars's Water, 452
Meteoroids from the Moon and Mars, 490

Mercury
Mercury, 460

Meteorites. *See* **Small Bodies**

Moons. *See* **Natural Planetary Satellites**

Natural Planetary Satellites
Callisto, 35
Enceladus, 204
Europa, 213
Ganymede, 244
Iapetus, 275
Io, 304
Jupiter's Satellites, 345
Lunar Craters, 360
Lunar History, 368
Lunar Interior, 374
Lunar Maria, 378
Lunar Regolith Samples, 384
Lunar Rocks, 389
Lunar Surface Experiments, 394
Mars's Satellites, 437
Miranda, 506
Neptune's Satellites, 535
Planetary Satellites, 617
Pluto and Charon, 641
Saturn's Satellites, 701
Titan, 852
Triton, 857
Uranus's Satellites, 901

The Neptunian System
Neptune's Atmosphere, 516
Neptune's Great Dark Spots, 520
Neptune's Interior, 524
Neptune's Magnetic Field, 527
Neptune's Ring System, 530
Neptune's Satellites, 535
Triton, 857

Planets and Planetology
Dwarf Planets, 85
Earth-Moon Relations, 90
Earth-Sun Relations, 94
Extrasolar Planets, 220
Extrasolar Planets: Detection Methods, 227
Impact Cratering, 279
Jovian Planets, 311
Jupiter's Atmosphere, 318
Jupiter's Great Red Spot, 324
Jupiter's Interior, 331
Jupiter's Magnetic Field and Radiation Belts, 335
Jupiter's Ring System, 341
Jupiter's Satellites, 345
Mars: Illusions and Conspiracies, 401
Mars: Possible Life, 405
Mars: Surface Experiments, 409
Mars's Atmosphere, 415
Mars's Craters, 421
Mars's Polar Caps, 428
Mars's Satellites, 437
Mars's Valleys, 441
Mars's Volcanoes, 446
Mars's Water, 452
Mercury, 460
Nemesis and Planet X, 510
Neptune's Atmosphere, 516
Neptune's Great Dark Spots, 520
Neptune's Interior, 524
Neptune's Magnetic Field, 527
Neptune's Ring System, 530
Neptune's Satellites, 535
Planetary Atmospheres, 564

Planetary Classifications, 572
Planetary Formation, 575
Planetary Interiors, 582
Planetary Magnetospheres, 589
Planetary Orbits, 594
Planetary Orbits: Couplings and Resonances, 601
Planetary Ring Systems, 605
Planetary Rotation, 612
Planetary Satellites, 617
Planetary Tectonics, 625
Planetology: Comparative, 630
Planetology: Venus, Earth, and Mars, 637
Pluto and Charon, 641
Saturn's Atmosphere, 674
Saturn's Interior, 683
Saturn's Magnetic Field, 686
Saturn's Ring System, 693
Saturn's Satellites, 701
Terrestrial Planets, 843
Uranus's Atmosphere, 886
Uranus's Interior, 890
Uranus's Magnetic Field, 893
Uranus's Rings, 896
Uranus's Satellites, 901
Uranus's Tilt, 906
Venus's Atmosphere, 914
Venus's Craters, 921
Venus's Surface Experiments, 925
Venus's Surface Features, 929
Venus's Volcanoes, 938

The Saturnian System
Enceladus, 204
Iapetus, 275
Saturn's Atmosphere, 674
Saturn's Interior, 683
Saturn's Magnetic Field, 686
Saturn's Ring System, 693
Saturn's Satellites, 701
Titan, 852

Scientific Methods
Archaeoastronomy, 1
Coordinate Systems, 65
Earth System Science, 99
Gravity Measurement, 252
Hertzsprung-Russell Diagram, 269
Infrared Astronomy, 286
Neutrino Astronomy, 538
Optical Astronomy, 558
Radio Astronomy, 659
Telescopes: Ground-Based, 824
Telescopes: Space-Based, 832
Ultraviolet Astronomy, 862
X-Ray and Gamma-Ray Astronomy, 947

Small Bodies
Asteroids, 6
Ceres, 40
Comet Halley, 44
Comet Shoemaker-Levy 9, 52
Comets, 56
Dwarf Planets, 85
Eris and Dysnomia, 208
Kuiper Belt, 352
Meteorites: Achondrites, 468
Meteorites: Carbonaceous Chondrites, 472
Meteorites: Chondrites, 477
Meteorites: Nickel-Irons, 481
Meteorites: Stony Irons, 486
Meteoroids from the Moon and Mars, 490
Meteors and Meteor Showers, 495
Oort Cloud, 555
Pluto and Charon, 641

The Solar System as a Whole
Earth's Origin, 173
Eclipses, 190
Interplanetary Environment, 295
Kuiper Belt, 352
Oort Cloud, 555
Solar System: Element Distribution, 767
Solar System: Origins, 773

The Stellar Context
Brown Dwarfs, 31
Gamma-Ray Bursters, 238
Main Sequence Stars, 398
Novae, Bursters, and X-Ray Sources, 543
Nuclear Synthesis in Stars, 549
Protostars, 649
Pulsars, 653
Red Dwarf Stars, 665
Red Giant Stars, 669
Stellar Evolution, 805
Supernovae, 816
Thermonuclear Reactions in Stars, 847
White and Black Dwarfs, 942

The Sun

Coronal Holes and Coronal Mass Ejections, 69
Earth-Sun Relations, 94
Solar Chromosphere, 713
Solar Corona, 717
Solar Evolution, 723
Solar Flares, 727
Solar Geodesy, 732
Solar Infrared Emissions, 736
Solar Interior, 739
Solar Magnetic Field, 743
Solar Photosphere, 748
Solar Radiation, 751
Solar Radio Emissions, 756
Solar Seismology, 759
Solar Structure and Energy, 763
Solar Ultraviolet Emissions, 778
Solar Variability, 784
Solar Wind, 787
Solar X-Ray Emissions, 792
Sunspots, 811

The Uranian System

Miranda, 506
Uranus's Atmosphere, 886
Uranus's Interior, 890
Uranus's Magnetic Field, 893
Uranus's Rings, 896
Uranus's Satellites, 901
Uranus's Tilt, 906

Venus

Venus's Atmosphere, 914
Venus's Craters, 921
Venus's Surface Experiments, 925
Venus's Surface Features, 929
Venus's Volcanoes, 938

A

Archaeoastronomy

Category: Scientific Methods

Archaeoastronomy is an interdisciplinary field studying how ancient civilizations sought answers to cosmological questions. Observations of sky phenomena influenced many aspects of ancient cultures, including but not limited to their science, religion, architecture, and lifestyles. Archaeoastronomy helps demonstrate how early technologies developed to better observe and understand the heavens. More often than not, the phenomena of greatest interest involved various solar system objects, especially the rising, setting, and apparent motions of the Sun, Moon, and bright planets.

OVERVIEW

Archaeoastronomy combines astronomy and archaeology into a study of ancient civilizations that focuses on the relationship between their observations of sky phenomena and their science, religion, architecture, and cultural practices. In advancing the discipline, astronomers and archaeologists work alongside ethnographers, geographers, anthropologists, mathematicians, historians, and others. However, many of the interpretations of archaeoastronomy are controversial and open to alternate explanations.

Archaeoastronomy dates back to the late 1600's and early 1700's, when the connection between ancient human-constructed structures and astronomical events was first made. Antiquarians like John Aubrey in 1678 and Henry Chauncy in 1700, studied the astronomical alignments of churches. Initial studies dealt with Middle Eastern and European cultures. Suggested by Euan MacKie, the term "archaeoastronomy" was first used in 1973 by Elizabeth Chesley Baity. Today, locations throughout the world are investigated.

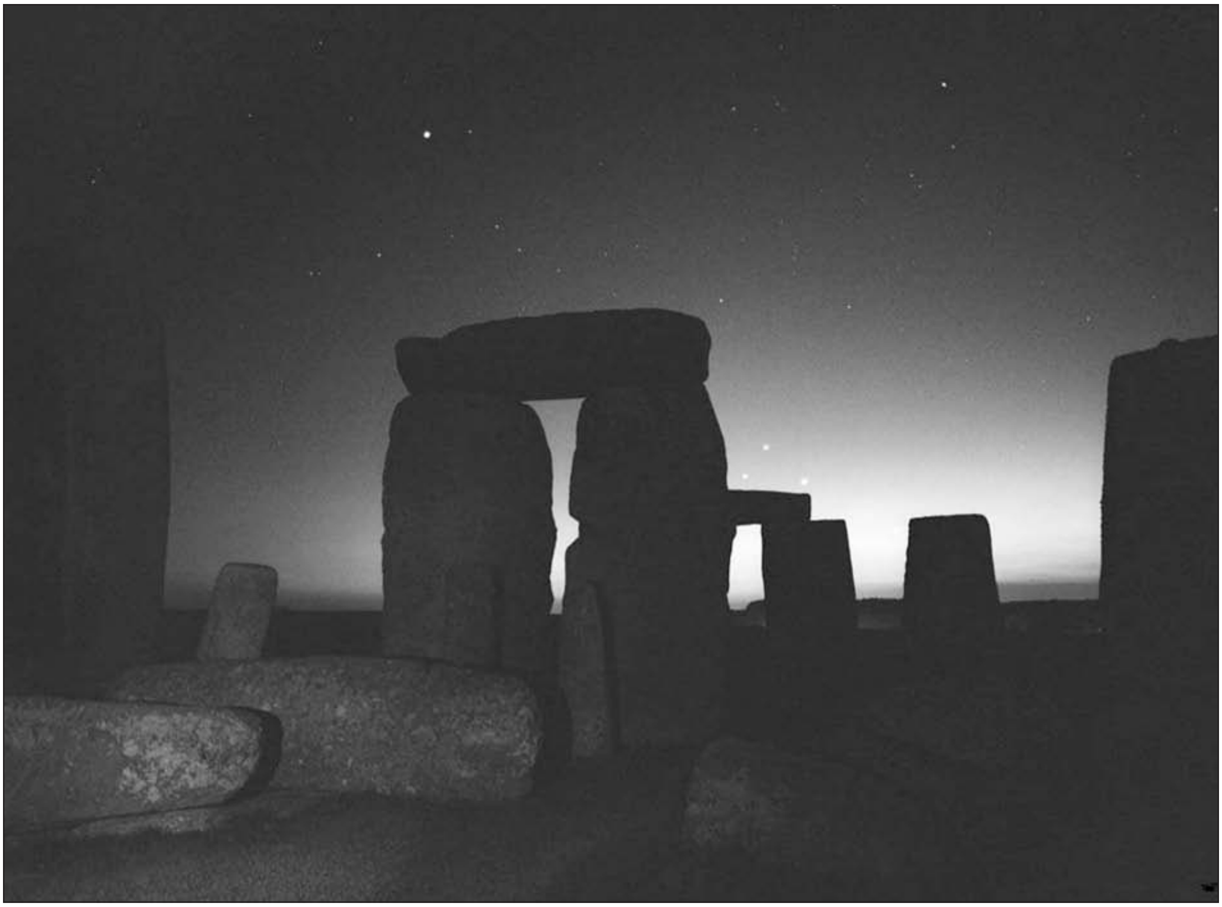
The best-known archaeoastronomy site is Stonehenge in Wiltshire, England. The current structure at the site has been dated to 2500 B.C.E. There is evidence of an older stone structure on the same site. An earthen bank and ditch that encompass the stones has been dated to 3100 B.C.E. The circular bank-and-ditch system measures 110 meters in diameter and has a large opening in the northeast and a smaller one to the south. Flint tools and bones of deer and oxen that were found in the ditch date the ring's age.

Large stones of the structure called sarsens were named for the quarry where they were mined, about 40 kilometers away from Stonehenge. Stones were carefully shaped before being positioned. Thirty were erected vertically in a 33-meter circle within the earthworks structure. Thirty lintel stones were placed on top of the sarsens. Stones are held together with tongue-and-groove joints. Each sarsen is about 4.1 meters tall and 2.1 meters wide, and each weighs about 25 tons. Lintel stones are slightly smaller. Inside this circle, five triliths were arranged in a horseshoe pattern with the open end facing northeast; each trilith is a structure composed of a lintel stone and two supporting vertical stones—that is, a structure consisting of three stones (hence the name, from "tri" for "three" and "lith" for "stone"). Triliths increased in size from 6 to 7.3 meters tall in the southwest. A single sarsen of the Great Trilith remains standing. The stone is 6.7 meters tall above ground level, with another 2.4 meters buried below ground.

In 1666, Aubrey was one of the first to study Stonehenge. His work was continued by William Stukeley, who incorrectly associated the site with the Druids. In 1740, John Wood created the most accurate map of Stonehenge. Some distance outside the sarsen circle is a pointed stone called the heel stone. When viewed from the center of the sarsen circle, the Sun rises directly over the heel stone during the summer solstice. The structure also is aligned with the winter solstice sunset and the southernmost moonrise. From outside Stonehenge, the winter sunset is framed by one particular trilith.

A second, wooden structure was built two miles away from Stonehenge around the same time. The Durrington Walls circle overlooks the Avon River and is remarkably similar to the stone structure. The timber circle (or timber henge), however, aligns with the winter sunrise. The avenue leading to the river aligns with the summer solstice sunset. Archaeoastronomer Mike Parker Pearson believes that the two structures were linked and that each played a key role in the culture and lives of their builders. The purpose of these two henges, however, is still unclear. The Stonehenge Riverside Project, led by Pearson, found evidence that Stonehenge was a burial site. Pearson believes that the site was used to bury thirty to forty generations of royal family members. The project also conducted excavations near the Durrington Walls site, finding at least three hundred homes believed to have been used as seasonal lodgings during celebrations held there.

The most famous Irish archaeoastronomy site is Newgrange. Built between 3300 and 2900 B.C.E., Newgrange is older than both the Giza pyramids (by five hundred years) and the Stonehenge triliths (by one thousand years). The dominant feature of this site is a 12-meter-high, 76-meter-wide mound that covers an acre of ground. The mound contains an 18-meter-long passage ending in a cross-shaped room. The site is believed to be a tomb. Cremated remains have been found inside the burial chamber. A wooden circle to the southeast and a smaller one to the south of the mound were added during the Neolithic period. The larger



Mercury, Venus, Mars, Jupiter, and Saturn can be seen in the night sky over Stonehenge. (Philip Perkins) circle has five rings of pits, with the outer one composed of wooden posts. The next inner ring was lined with clay and was used to cremate animals buried in the inner three circles of pits. The mound was surrounded by a circle of large, freestanding stones added during the Bronze Age. Newgrange was not discovered until the late seventeenth century because of slippage covering the mound's entrance. Michael O'Kelley and his team excavated and restored the site between 1962 and 1975.

The complexity of the mound's construction becomes obvious on the winter solstice, when sunlight enters the tomb through a roof box above the entrance and shines directly down the passageway onto the chamber floor. Solar alignments have been found in other tombs, but none is as precise as Newgrange, and none includes a roof box.

The only remaining Great Wonder of the ancient world is the Great Pyramid of Giza, located near Cairo, Egypt. That pyramid is also known as the Pyramid of Khufu or the Cheops Pyramid. The base of the structure is nearly a perfect square, each side measuring 230 meters. The Great Pyramid is almost fifty stories high, measuring 146.7 meters. Two slightly smaller pyramids (the Pyramid of Khafre and the Pyramid of Menkaure) also occupy the Giza complex. The three pyramids all follow north-south and east-west lines within a fraction of a degree. Archaeoastronomers propose various explanations for that orientation. In 2000, Egyptologist Kate Spence published an article with a new theory for the age and alignment of the pyramids. Spence believes that pyramid construction began between 2485 and 2475 B.C.E. Her theory also explains why all the pyramids are slightly offline with true north. Spence argues that the builders used two faint stars, Kochab (in the Little Dipper) and Mizar (in the Big Dipper), as guide stars. In 2467 B.C.E. these two stars were aligned directly with north and the celestial pole star. Egyptian astronomers could locate the two stars and use a plumb line to find the northern horizon. Pyramid construction took so long that Earth's precession eventually caused Kochan and Mizar to lose their exact alignment with north. This theory explains the orientation of the pyramids and gives a more exact estimate of their age. Other scientists argue that the three pyramids simply represent the three stars in Orion's belt.

The most famous Mayan city is Chichén Itzá in Mexico. At the center of the city lies El Castillo, a temple to the god Kukulcan (Quetzalcoatl). Built between 1000 and 1200 C.E., El Castillo is a step pyramid with staircases on each side leading up to a central platform at the top. Each side has 91 steps. Counting the top platform, in total there are 365 steps, one for each day in the Haab, a part of the Mayan calendar. Sculptures of serpents adorn the side of the north-facing staircase. At the vernal and autumnal equinoxes, they cast a shadow, making it appear as though a serpent is climbing down the stairs. The western side is also aligned with the setting Sun around May 25, the day that usually marks the transition between dry and rainy seasons.

Another important Mayan site is Uxmal, located on the Yucatan peninsula. The Governor's Palace is a long, low structure built atop a large platform. The palace is aligned with the southernmost rising of Venus, which occurs once every eight years. The temple is also aligned with the pyramid of Cehtzuc on an azimuth of 118°. Standing in the entrance of the palace, Venus's southernmost rising would be seen directly over the Cehtzuc pyramid 4.6 kilometers away. Venus was very important to the Maya. The Governor's Palace is covered with hundreds of glyphs representing Venus, one of the most important Mayan deities. The Maya built observatories and kept very accurate records of Venus. They recognized that Venus alternated between appearing for several months in the eastern sky before sunrise and then for several months in the western sky after sunset. The Maya used these observations and those of other celestial objects to create one of the most accurate ancient calendars.

Medicine wheels are circular stone structures, with "spokes" radiating outward from a central cairn (pile of stones). The wheels were given the name "medicine" from the speculation that they were used by Native Americans in healing rituals. In 1885, when George Dawson published the first paper about medicine wheels, more than 20,000 of these artifacts had

Built between 1000 and 1200 C.E., El Castillo is a step pyramid with staircases on each side leading up to a central platform at the top. There are 365 steps, one for each day in the Haab, a part of the Mayan calendar. Sculptures of serpents adorn the side of the north-facing staircase. At the vernal and autumnal equinoxes, the shadows they cast make it appear as though a serpent is climbing down the stairs. The western side is aligned with the setting Sun around May 25, marking the transition between dry and rainy seasons. (©Pierdelune/Dreamstime.com)

been identified throughout North America. Today, only 135 survive in the United States and Canada. Some wheels are more than forty-five hundred years old, while others were built in later centuries. Medicine wheels vary in type. Some have spokes that extend out of the circle up to 120 meters. Others contain more rings of stones. The origin and true purpose of medicine wheels will never be known conclusively, but they are bear an unmistakable connection to the night and daytime skies.

The southernmost surviving ring is Bighorn Medicine Wheel, high in the Bighorn Mountains of north-central Wyoming, a site tentatively dated between 1200 and 1700 C.E. In the 1970's, John A. Eddy studied medicine wheels and believes that they were astronomically aligned. The Bighorn wheel has twenty-eight spokes radiating from a central cairn, each spoke ending in its own cairn. The number of spokes is the same as the number of days in a lunar cycle. One of the spokes extends out of the ring by 3.96 meters. Eddy discovered that the longer spoke aligned with sunrise on the summer solstice. Further support for his theory is the fact that Bighorn Medicine Wheel, at an elevation of 9,956 feet, is almost constantly under snow, with the exception of two months in the summer around the solstice. Another spoke aligns with the solstice sunset. Eddy linked other cairns with alignments to the rising of the stars Sirius (in Canis Major), Aldebaran (in Taurus), and Rigel (in Orion). Since these stars hold significance for several Native American tribes, Eddy believes that the wheel was used as a calendar or timepiece.

Eddy and other scientists have found similar alignments with other medicine wheels. Opponents to Eddy's theory argue that Native Americans used medicine wheels for religious or spiritual ceremonies. Having no evidence other than the wheels themselves, archaeoastronomers may never solve all of their mysteries.

METHODS OF STUDY

There are two main methodologies used by archaeoastronomers. Green archaeoastronomy relies only on alignments and is used when little or nothing is known about the builders of a site. Some scientists criticize this approach because it postulates that the ancients were interested in astronomy but offers no insights into why or what role it played in their cultures. Brown archaeoastronomy, by contrast, is often compared with cultural history or the history of astronomy. Brown archaeoastronomy studies historical and ethnographic records to learn about the roles astronomy played in ancient cultures. (The names of these two methods have no import other than their reference to the colors of the covers of the first published texts proposing and explaining the nature of each technique; green came first, followed by brown.)

At some sites, both methods are used. Alexander Thom developed many of the methods used in green archaeoastronomy. He examined British sites, looking for evidence of astronomy in the builders' society. Thom looked for alignments with the Sun (on solstices and equinoxes), the Moon, and the stars. One significant alignment, he believed, could be the result of chance, but more than one, he postulated, indicated deliberate attention to celestrial objects on the part of the ancient builders. Euan MacKie, one of Thom's strongest supporters, studied British and Mayan sites, looking for proof of a link between the two societies.

Scientists used brown archaeoastronomy to interpret the Mayan city of Chichén Itzá. They studied Mayan records trying to determine what astronomical bodies were important in their culture. Once archaeoastronomers learned the importance of Venus to the Maya, they evaluated the site and found that many buildings were aligned with the rising or setting of the planet.

Anthony Aveni uses the Incan capital Cuzco as an example of the value of brown methods. Records of Incan life and legends were written by the Spanish and provide the majority of information known about the culture. These records explain that Cuzco was planned extensively before it was constructed. Most structures have astronomical alignments—with sunrise or sunset on the solstices, for example.

CONTEXT

Although controversial, archaeoastronomy—the study of the beliefs, religions, science, architecture, and cosmology of ancient cultures—is an increasingly important scientific field. The astronomical alignments of buildings, structures, and cities not only provide a new way of thinking about ancient societies but also elucidate what was known about the solar system before modern astronomy arose in the sixteenth century. The organizational, engineering, scientific, and mathematical skills of these cultures are showcased by their archaeoastronomy sites. Archaeoastronomy also expands the study of the history of science and religion further into the past than has been possible by the study of written records.

However, many questions remain unanswered. What caused societies in Russia, China, Egypt, Spain, and in other parts of the world to build pyramid-shaped structures? What led these cultures to study and worship the heavens? The study of archaeoastronomy can also reveal some things about ourselves. Many people today pay no attention to the night sky. Most cannot explain what an eclipse or solstice is. Interest in the exploration of the solar system has declined since the 1960's even as the technology to discover more about the solar system and the universe beyond has expanded greatly. Less than one-third of all Americans and Europeans can see the Milky Way from their own backyards because of urban light pollution. Today's societies lack a connection with the heavens, something that was a main component of ancient cultures. One can only wonder how the builders of Stonehenge or the Mayan temples would view this modern attitude.

Jennifer L. Campbell

FURTHER READING

Aveni, Anthony. *Stairways to the Stars: Skywatching in Three Great Ancient Cultures*. New York: Wiley, 1997. An informative work focusing on the Mayan Great Pyramid, the Incan capital Cuzco, and Stonehenge. Aveni speculates about how these ancient structures were built and discusses the various theories for their usage and origins.

Chamberlain, Von Del, John B. Carlson, and M. Jane Young. *Songs from the Sky: Indigenous Astronomical and Cosmological Traditions of the World*. College Park, Md.: Ocarina Books, 2005. A collection of papers written by scientists from various fields dealing with cultures of the early Americas and the Near East. Astronomy and cosmology in these societies are investigated, including their roles in architecture.

Fountain, John, and Rolf Sinclair. *Current Studies in Archaeoastronomy: Conversations Across Time and Space*. Durham, N.C.: Carolina Academic Press, 2005. A collection of papers from the fifth Oxford Conference on Archaeoastronomy focusing on sociology, astronomy, and culture of various ancient civilizations.

Kelley, David, and Eugene Milone. *Exploring Ancient Skies: An Encyclopedia Survey of Archaeoastronomy*. New York: Springer, 2004. A comprehensive look at archaeoastronomy, focusing on observational astronomy and its history. Examines the ways ancient cultures watched celestial objects and used them to keep time and develop calendars. Discusses origins of astronomy in cultures throughout the world, including Mesoamerica, Egypt, Africa, China, Korea, India, Japan, Greece, and Europe.

Magli, Giulio. *Mysteries and Discoveries of Archaeoastronomy: From Giza to Easter Island*. New York: Springer, 2009. A comprehensive study of all areas of modern archaeoastronomy. Takes the reader on a tour of the world's greatest ancient sites and discusses their astronomical alignments and cultural astronomy. Asserts that these great structures were built as symbols and the foundation of power. Discusses the Giza pyramid complex in great detail and hypothesizes that the Cheops and Chephren pyramids were built as part of the same structure.

See also: Coordinate Systems; Earth System Science; Optical Astronomy; Planetary Orbits; Telescopes: Ground-Based.

Asteroids

Category: Small Bodies

Asteroids are minor bodies of a wide variety of sizes in orbit around the Sun, primarily but not exclusively located between the orbits of Mars and Jupiter. They provide important clues regarding the early history of the solar system, including the effect of their collisions on the surfaces of planets or their satellites. A class popularly referred to as near-Earth asteroids threaten to impact Earth.

OVERVIEW

Although discovery of the first asteroid was accidental, it came as no surprise to the astronomical community of the day. In 1766, German astronomer Johann Titius (1729-1796) observed that the positions of the planets could be approximated very closely by a simple empirical rule. Adding 4 to each number in the sequence {0, 3, 6, 12, 24, 48 . . . } and dividing the sum by 10 yields the mean planetary distances from the Sun in astronomical units (the distance from the Earth to the Sun is one astronomical unit, or 1 AU). The exception to this rule is the fifth element in that purely mathematical sequence, where an apparent gap occurs at 2.8 AU. It must be noted that this is just an empirical observation with no known physical basis. This rule was publicized by Johann Bode (1747-1826) and led to a search for a missing planet in the gap between Mars at 1.5 AU and Jupiter at 5.2 AU.

On January 1, 1801, the Sicilian astronomer-monk Giuseppe Piazzi (1746-1826) accidentally discovered a moving object during a routine star survey. He named it Ceres, for the patron goddess of Sicily. Soon its orbit was calculated by Carl Friedrich Gauss (1777-1855). At 2.77 AU, Ceres was found, coincidentally, to conform closely to the Titius-Bode rule. However, since Ceres seemed to be too small to be classified as a planet, the search continued.

In March, 1802, German astronomer Heinrich Olbers (1758-1840) found a second minor body at the same predicted distance. He named it Pallas. In 1803, Olbers proposed that meteorites come from an exploded planet near 2.8 AU. This possibility led to a continued search resulting in the discovery of Juno in 1804 and Vesta in 1807. The latter discovery again was made by Olbers. It took quite some time for a fifth small body to be discovered (in 1845), but by 1890 the total had reached three hundred. These bodies came to be called "asteroids" for their faint, starlike images. In 1891, German astronomer Max Wolf (1863-1932) began using a long-exposure camera to detect asteroids. Since then, thousands of asteroids have been registered in the official catalog of the Institute of Theoretical Astronomy in Leningrad.

Asteroids are usually referred to officially by both a number and a name, such as 3 Juno or 1,000 Piazzi. About one hundred newly numbered asteroids are cataloged each year. Sky surveys indicate as many as 500,000 asteroids that are large enough to appear on telescopic photographs. Most asteroids are found within the main asteroid belt, which extends from 2.1 to 3.4 AU, and about half are between 2.75 and 2.85 AU. Asteroids revolve around the Sun in the same direction as the planets but tend to have more elliptical orbits. Their orbits are inclined up to 30° to the ecliptic plane, but they are far less eccentric than comet orbits. The smallest asteroids are a few kilometers wide; the largest, 1 Ceres (now considered a "dwarf planet"), about 1,000 kilometers wide. In 1867, American astronomer Daniel Kirkwood (1814-1895) discovered gaps in the asteroid belt where relatively few asteroids are found. These so-called Kirkwood gaps occur where asteroids have orbital periods that are simple fractions of the twelve-year revolution period of the giant planet Jupiter about the Sun, resulting in periodic gravitational influences called resonances. Such depletions occur, for example, at about 3.3 AU (where the periods have a six-year, 1:2 resonance with Jupiter) and 2.5 AU (a four-year, 1:3 resonance); other resonances, however, act to stabilize certain asteroids, such as the Hilda

An artist's rendering of one of two asteroid belts of Epsilon Eridani, a solar system located in the constellation Eridanus. (NASA/JPL-Caltech)

Facts About Selected Asteroids

Name	*Diameter (km) or Dimensions*	*Mass (10^{15}kg)*	*Rotational Period (hrs)*	*Orbital Period (yrs)*	*Distance from Sun (AU)*
Ceres (dwarf planet)	960 × 932	870,000	9.075	4.60	2.767
Palas	570 × 525 × 482	318,000	7.811	4.61	2.774
Juno	240	20,000	7.210	4.36	2.669
Vesta	530	300,000	5.342	3.63	2.362
Eugenia	214	6,100	5.699	4.49	2.721
Siwa	110	1,500	18.5	4.52	2.734
Ida	58 × 23	100	4.633	4.84	2.861
Mathilde	66 × 48 × 46	103.3	417.7	4.31	2.646
Eros	33 × 13 × 13	7.2	5.270	1.76	1.458
Gaspra	19 × 12 × 11	10	7.042	3.29	2.209
Icarus	1.4	0.001	2.273	1.12	1.078
Geographos	2.0	0.004	5.222	1.39	1.245
Apollo	1.6	0.002	3.063	1.81	1.471
Chiron	180	4000	5.9	50.70	13.633
Shipka	—	—	—	5.25	3.019
Rodari	—	—	—	3.25	2.194
McAuliffe	2-5	—	—	2.57	1.879
Mimistrobell	—	—	—	3.38	2.249
Toutatis	4.6 × 2.4 × 1.9	0.05	130	3.98	2.512
Nereus	2	—	—	1.82	1.490
Castalia	1.8 × 0.8	0.0005	—	1.10	1.063
Otawara	5.5	0.2	—	3.19	2.168
Braille	2.2 × 1.0	—	—	3.58	2.341

Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.

group at 4 AU (2:3 resonance), which is named for 153 Hilda.

Some asteroids have orbits departing greatly from the main belt. In 1772, French mathematician Joseph Lagrange (1736-1813) showed that points in Jupiter's orbit 60° ahead of and behind the planet are gravitationally stable (1:1 resonance). In 1906, Max Wolf discovered the first so-called Trojan asteroid, 588 Achilles, at the Lagrangian point 60° ahead of Jupiter. Subsequent discoveries have revealed several hundred Trojan asteroids. Those ahead of Jupiter are named for Greek heroes, and those behind are named for Trojan heroes; there is one Greek spy (617 Patroclus asteroid) in the Trojan group, and one Trojan spy (624 Hektor) in the Greek group. Hektor is the largest known Trojan asteroid, at about 150 by 300 kilometers, and is the most elongated of the more massive asteroids. At least two objects have orbits that extend beyond Jupiter: 944 Hidalgo, which may be a burned-out cometary nucleus, and 2060 Chiron, whose orbit extends beyond Saturn.

Some asteroids depart from the main belt over only part of their orbit. Mars-crossing Amor group bodies have elongated orbits that carry them inside Mars's orbit but still keep them well outside Earth's orbit. The Martian satellites Phobos and Deimos have long been suspected by many to be captured asteroids, perhaps from this group. Apollo group members come inside Earth's orbit. (The groups were named for their first examples, discovered in 1932.) Estimates indicate about thirteen hundred Apollos ranging in from 0.4 to 10 kilometers across, with an estimated average Earth-collision rate of about one in 250,000 years. The closest known approaches were Hermes, in

1937, at about 780,000 kilometers, and 1566 Icarus, in 1968, at about 6 million kilometers. Smaller Apollos may be an important source of meteorites, and 100-meter objects capable of making a 1-kilometer crater strike Earth about every two thousand years. Aten-type asteroids are Earth-crossers with elliptical orbits smaller than Earth's. Some asteroids appear to be grouped in families that may be the fragments resulting from an earlier collision between asteroids.

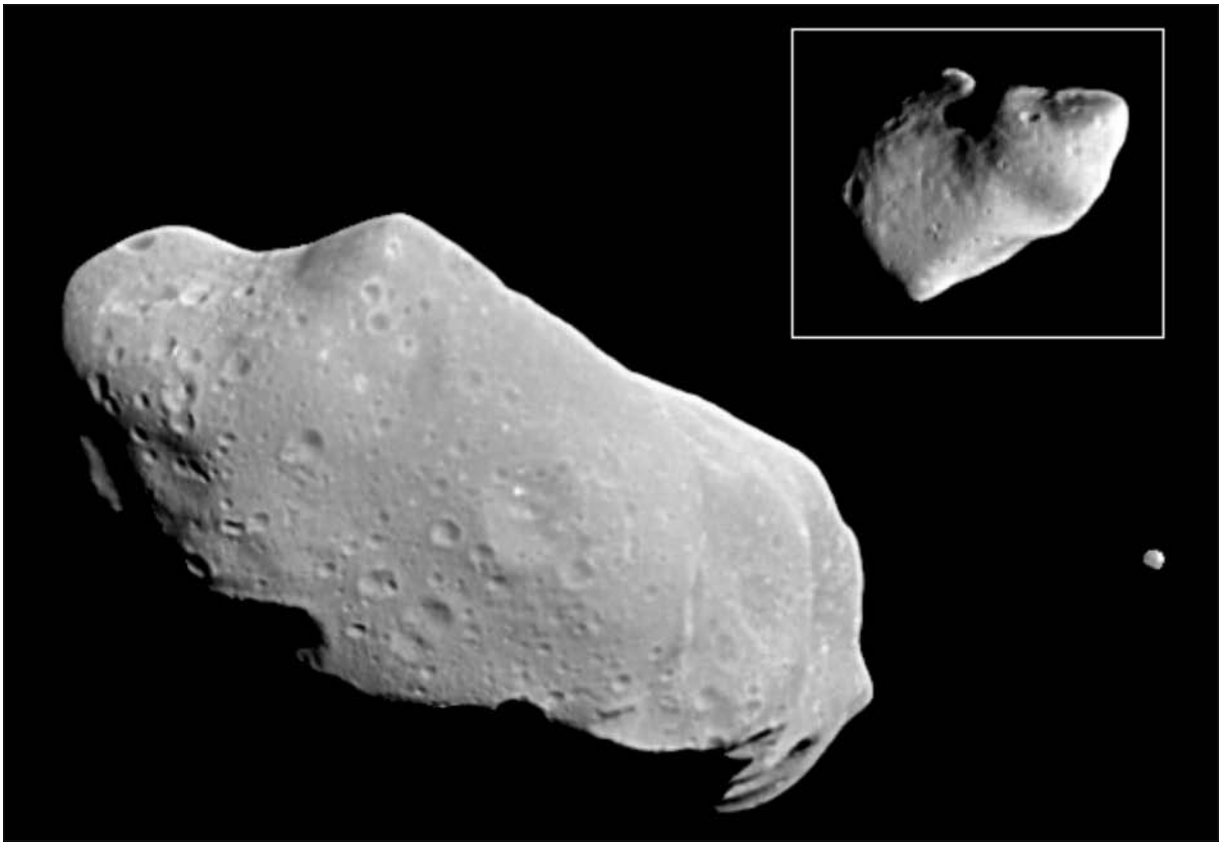
Chemical and physical characteristics of asteroids are mostly determined by remote-sensing techniques that study electromagnetic radiation reflected off their surfaces. More than five hundred asteroids have been studied by remote sensing and radar astronomical techniques. These studies have indicated asteroidal compositions similar to those of meteorites. Comparison with reflected light from meteorites suggests several classes. Rare E-type asteroids possess the highest albedo (23 to 45 percent reflection). They appear to be related to enstatite (a magnesium silicate mineral) chondrites, and are concentrated near the inner edge of the main belt. About 10 percent of asteroids are S-type; they have relatively high albedos (7 to 23 percent) and appear reddish in color. They likely are related to stony chondrites, are found in the inner to central regions of the main belt, and generally range in size from 100 to 200 kilometers. The largest S-type is 3 Juno, at about 250 kilometers in diameter. Much smaller Apollo asteroids are also in this category. A few asteroids in the middle belt are classified as M-type, since their reflected light (7 to 20 percent) reveals evidence of large amounts of nickel-iron metals on their surface, similar to iron or stony-iron meteorites.

About three-quarters of all asteroids are C-type, having relatively low albedos (2 to 7 percent) and grayish colors similar to that of the Earth's moon. They are found in the outer belt and among the Trojans. They resemble carbonaceous chondrite meteorites, containing water-bearing silicate-based and carbon-based minerals along with some organic compounds (about 1 percent). The largest of all the asteroids, 1 Ceres (now considered a dwarf planet), is in this category. Some evidence supports the claim that Ceres has a mixture of ice and carbonaceous minerals on its surface. Dark reddish, D-type asteroids are found in the same regions and have similar albedos.

About 10 percent of asteroids remain unclassified and are designated as U-type. In general, asteroids with low-temperature volatile materials lie farther from the Sun, whereas those in the inner part of the main belt are richer in high-temperature minerals, displaying little evidence of volatile water and carbon compounds.

Many asteroids exhibit periodic variations in brightness that suggest irregular shapes and rotation. Their measured rotational periods range from about three to thirty hours. There is some evidence that S-type asteroids rotate faster than C-type asteroids but more slowly than M-type asteroids. Large asteroids (greater than 120 kilometers) rotate more slowly with

Five of the Largest Asteroids Shown Relative to the Western U.S.



Asteroid Ida, so named for its potato shape, as imaged by the Galileo spacecraft. Ida's satellite Dactyl appears in the inset. (NASA/JPL/USGS)

increasing size, but small asteroids rotate more slowly with decreasing size, suggesting that large asteroids may be primordial bodies, while smaller ones may be fragments produced by collisions. Calculations show that rotation rates longer than two hours produce centripetal forces weaker than gravity, which indicates that loose debris can exist on the surface of even the fastest known rotating asteroid, the Apollo object 1566 Icarus, which has a 2.25-hour rotation rate. Polarization studies of light reflected from asteroids indicate that many do have dusty surfaces.

Named after the Greek god of love, Eros is an S-type asteroid belonging to the Amor group. As the second-largest of the near-Earth asteroids, it is larger than the asteroid generally accepted to have been responsible for the extinction of the dinosaurs that impacted Earth near the Yucatán peninsula 65 million years ago. Eros is 13 by 13 by 33 kilometers in size. It was the first asteroid recognized to approach inside the orbit of Mars, and in 1975 it became the first asteroid to be studied with Earth-based radars.

Computer models suggest the possibility that larger asteroids have a deep layer of dust and rock fragments (or regolith) similar to that on the surface of the Moon. Asteroids with diameters larger than 100 kilometers are believed to have undergone a process of differentiation in which heavier metals sank to the core, leaving a stony surface of lighter materials later pulverized by collisions to form a layer of dust.

Asteroid elongations can be estimated from the change in brightness, which can vary by a factor of three or more. For example, radar evidence indicates that the unusual Trojan asteroid 624 Hektor (150 by 300 kilometers) may be a dumbbell-shaped double asteroid. Kilometer-scale asteroids have been observed with lengths up to six times greater than their width. Main-belt asteroids tend to be less elongated than Mars-crossers of the same size, perhaps because of greater erosion from collisions in the belt. Asteroids larger than about 400 kilometers tend to be more spherical, since their gravitational attractions exceed the strength of their rocky materials, causing deformation and plastic flow into a more symmetric shape.

An asteroid's size occasionally can be determined quite accurately by timing its passage in front of a star, that is, in a stellar occultation. In a few cases, stellar light has been occulted more than once in a single passage, indicating that asteroids may possess satellites. Radar-based studies have confirmed this theory. Also, as the Galileo spacecraft flew through the main belt on its way to enter orbit in the Jupiter system, it imaged a satellite revolving about an asteroid. The irregularly shaped asteroid Ida was discovered to have a small satellite later named Dactyl. Ida is a member of the Koronis family and of S-type, which is 56 by 24 by 21 kilometers in size and rotates once around its own axis every 278 minutes. Dactyl is only 1.2 by 1.4 by 1.6 kilometers in size and is also of S-type. This strongly suggests that it was created when a larger asteroid smashed into Ida. Previously, the Galileo spacecraft had also provided the first close-up images of an asteroid, when it passed within five thousand kilometers of the 19- by 12- by 11-kilometer-sized S-type asteroid Gaspra on October 29, 1991. Gaspra has an irregular shape, one resembling a potato.

The distribution of asteroid sizes and masses supports the idea that many have undergone a process of fragmentation. Typical relative velocities of encounter, about 5 kilometers per second in the main belt, are quite adequate to fragment most asteroids. Ceres contains nearly half the mass of all the asteroids, but it is more than three times smaller than the Moon and about fifty times less massive. About 80 percent of the total mass of all asteroids is contained in the four largest ones, and only about ten are larger than 300 kilometers. Studies suggest that the main belt was several times more massive in the past but that in the process of fragmentation, the smallest dust particles were removed by radiation pressure from the Sun.

Interest in asteroids increased when strong evidence was advanced to solve the mystery of the demise of the dinosaurs 65 million years ago. Physicist Luis Alvarez and his geologist son Walter sampled the worldwide clay layer that marks the end of the Cretaceous period and the start of the Tertiary period (the so-called K-T boundary, which essentially marks the demarcation between the age of dinosaurs and the rise of mammals within the fossil record). This thin layer of clay is enriched in the rare elements of iridium and osmium, having levels more akin to asteroids than Earthly materials. Thus, the impact theory for killing off the dinosaurs was proposed, and largely accepted except by certain portions of the paleontology community. That is, until a crater dated to 65 million years was discovered off the coast of the Yucatán peninsula. Some still insist that more than an asteroid impact was necessary to account for the observed diminishment of dinosaur species leading up to the extinction event 65 million years ago. However, the majority of the scientific community has come to accept the asteroid impact theory, at least as the principal cause of the sudden mass extinction at the end of the Cretaceous period. Since this event marks the boundary between the Cretaceous and Tertiary periods, it is often referred to as the K-T event.

This spurred interest in asteroid and comet impacts causing extreme environmental damage to the Earth at other times in the past, along with a desire to search for near-Earth asteroids that might represent a threat in the future. Twenty-five years after the proposal that an asteroid impact killed the dinosaurs received initial lukewarm acceptance by paleontologists; some researchers proposed that an even bigger asteroid (or comet) impact was responsible for the so-called Great Dying, the mass extinction at the end of the Permian period that closed out the Paleozoic era. At the end of the Permian 248 million years ago, more than 95 percent of all species died off rather suddenly; life nearly did not make it into the Mesozoic era, during which the dinosaurs eventually arose to dominance.

Researchers point to a large crater in the Antarctic (1.5 kilometers under the ice pack that dates to the time of the Permian mass-extinction event) as well as heavily jumbled areas in Siberia (known as the Siberian Traps), that might have received tremendous seismic energy after impact energy would have undergone antipodal focusing off Earth's core. The Siberian Traps also was an area of tremendous volcanic activity at the end of the Permian period. Was this coincidental or the result of an impact with antipodal focusing of seismic energy? In 2008 this theory remained highly speculative, rather than enjoying the widespread acceptance of the K-T event that killed the dinosaurs. However, if the theory is correct, such an event underscores the danger posed by asteroid and comet impacts on Earth.

Impact of even a small asteroid could pose a tremendous threat to human civilization. Throughout the 1990's and the early twenty-first century, a number of newly discovered near-Earth asteroids were thought to have a significant chance of hitting Earth in the quite near future. However, in each of those cases, additional observations refined the asteroid's orbit to the point where it was clear it would not hit Earth after all. There remained one major exception, however. Discovery of the asteroid 99942 Apophis, a member of the Aten group, led to major concern beginning in late 2004 that this 350-meter-across rock had a relatively worrisome potential to impact Earth in 2029. Precise observations of Apophis's orbit, ranging from 0.746 AU to 1.099 AU, dramatically lowered the probability that it would strike the Earth. However, Apophis would indeed pass within the altitude of geosynchronous satellites, less than 36,000 kilometers from Earth's surface. If Apophis passed within a special corridor only 400 meters across, gravitational influences could cause it to return and strike the Earth on Friday, April 13, 2036.

The Torino scale assesses the relative impact hazard of an asteroid impact. For a time after its discovery, Apophis rated a level 4 on the Torino scale; that is the highest level of threat. Further orbit refinements lowered the threat assessment to a level 0 threat, but, after realization of the possible return in 2036, it was raised to Level 1. Although Apophis will come very close to Earth in 2029, refinement of available orbital data has since determined that the chances of Apophis hitting the Earth in 2036 are more comforting: less than 1 in 45,000. The 2036 encounter will set up another close encounter the following year, but the chances that this would result in an Earth impact are calculated to be less than 1 in 12.3 million. Nevertheless, Apophis points out the absolute requirement for close monitoring of asteroids, particularly the near-Earth ones, and the development of means whereby asteroids could be deflected or destroyed in order to preserve Earth's biosphere and save human civilization. This sort of natural megadisaster is one of the few that humans have the potential to mitigate or prevent if action is taken sufficiently early once the threat is identified.

Hollywood has even taken notice of the asteroid or comet impact threat to Earth. Several scientifically incorrect action movies were produced, some of which were popularly received. Many of these movies, such as *Armageddon* (1998), portray the use of some type of nuclear device as the only viable way to avert an asteroid impact. In many real cases, nuclear explosions detonated within, on the surface of, or close to asteroids either would be insufficient or could merely fragment it so badly that an even worse situation, a swarm of impacting bodies, might ensue.

METHODS OF STUDY

Studies of asteroids hold the potential for expanding our understanding of the formation of bodies of sizes between the smallest objects and full planets, and also could lead to development of technology to prevent an impact that might devastate life on Earth and even wipe out civilization.

The Galileo spacecraft passed near enough to two asteroids to photograph them directly. The NEAR spacecraft orbited Eros and later landed on its surface, providing close-up photographs of an asteroidal surface. For the most part, however, indirect methods of remote sensing must be used to determine asteroidal properties by studying the reflected electromagnetic radiation that comes from their surfaces. These methods include photometry, infrared radiometry, colorimetry, spectroscopy, polarimetry, and radar detection. They can be augmented by

The Torino Asteroid Impact Hazard Scale

Scale	*Description*
0	Event having no likely consequences: The likelihood of a collision is zero, or low enough to be effectively zero. This designation also applies to any small object that, in the event of a collision, is unlikely to reach the Earth's surface intact.
1	Event meriting careful monitoring: The chance of collision is extremely unlikely, about the same as a random object of the same size striking the Earth within the next few decades.
2	Event meriting concern: A somewhat close but not unusual encounter. Collision is very unlikely.
3	Event meriting concern: A close encounter, with 1 percent or greater chance of a collision capable of causing localized destruction.
4	Event meriting concern: A close encounter, with 1 percent or greater chance of a collision capable of causing regional devastation.
5	Threatening event: A close encounter, with a significant threat of a collision capable of causing regional devastation.
6	Threatening event: A close encounter, with a significant threat of a collision capable of causing a global catastrophe.
7	Threatening event: A close encounter, with an extremely significant threat of a collision capable of causing a global catastrophe.
8	Certain collision: A collision capable of causing localized destruction. Such an event occurs somewhere on Earth between once per 50 years and once per 1,000 years.
9	Certain collision: A collision capable of causing regional devastation. Such an event occurs between once per 1,000 years and once per 100,000 years.
10	Certain collision: A collision capable of causing a global climatic catastrophe. Such an event occurs once per 100,000 years or less often.

comparative studies with meteorites, whose composition and structure can be analyzed by direct methods in the laboratory. Such methods include chemical, spectroscopic, and microscopic analysis, and processes of fragmentation can be studied by producing high-speed collisions between comparable materials in the laboratory. Such comparative studies must recognize various differences between meteorites and asteroids. The masses of only the three largest asteroids have been determined from their gravitational effects on other bodies; their densities are between 2.3 and 3.3 grams per cubic centimeter.

Photometry is the study of how light is scattered by various surfaces. The varying brightness of reflected sunlight from asteroids can be measured by photoelectric observations to determine their rotation periods and approximate shapes. One test of this method was made in 1931, when the Amor asteroid 433 Eros came close enough (23 million kilometers) for scientists to observe the tumbling motion of this elongated object (7 by 19 by 30 kilometers) and to confirm its 5.3-hour rotation. The size of an asteroid can be estimated from its brightness together with its distance, orbital position, and albedo. The albedo is important, since a bright, small object may reflect as much light as a dark, large object. Since a dark object absorbs more heat than a light object, albedos can be determined by comparing reflected light with thermal radiation measured by infrared radiometry. Photometric measurements also give information on surface textures. Colorimetry involves measuring the range of wavelengths in the reflected light to determine surface colors. Most asteroids are either fairly bright, reddish objects (with albedos of up to 23 percent) composed largely of silicate-type materials or grayish objects, at least as dark as the Moon (11 percent albedo), composed of carbonaceous materials.

Spectroscopy is the spectral analysis of light and can be used to infer the composition of many asteroids. Optical and infrared reflectance spectra exhibit absorption bands at characteristic frequencies for given materials. An asteroid's surface composition is determined by comparing its spectrum with the spectra of light reflected from meteorites of known composition. Examples of this method applied to U-type (unclassified) asteroids include the identification of the silicate mineral pyroxene in the infrared spectrum of Apollo asteroid 1685 Toro, and the matching of the surface of Vesta with a basaltic achondrite that resembles lava. Most asteroids appear to have unmelted surfaces with little or no evidence of lava eruptions. About two-thirds of the Trojans are D-type asteroids with no known meteorite counterparts because of their distance from Earth. Their spectra have been matched with the spectra of coal-tar residues, suggesting possible organic compounds.

Polarimetry uses measurements of the alignment of electric field vibrations of the reflected sunlight and its variation with direction to estimate albedos. Polarization measurements have also been interpreted as evidence for dust-covered surfaces, but they leave uncertainty about the depth of the dust layer. Radar observations of Eros during a close approach to Earth in 1975 were made at a wavelength of 3.8 centimeters and indicated that the surface must be rough on a scale of centimeters. Since optical polarimetry suggests that Eros is dusty, the radar results imply that the dust must be too thin to smooth rock outcrops of more than a few centimeters. Radar measurements also provided independent estimates of the size of Eros, confirming photometric estimates of its dimensions. The NEAR spacecraft confirmed these observations.

As spacecraft results such as this demonstrate, the best method to study asteroids is by means of a space probe. When Pioneers 10 and 11 passed through the asteroid belt, scientists found that it has no more dust than any other part of the solar system. The Galileo probe encountered Gaspra in 1991 and Ida in 1993, both S-type asteroids. The probe determined the masses, sizes, and shapes. The Cassini spacecraft on its way toward orbit about Saturn flew through the asteroid belt and passed asteroid 2685 Masursky at a distance of 1.6 million kilometers. Named after the famed planetary scientist Hal Masursky, this body was a little understood 15- by 20-kilometer asteroid prior to Cassini's encounter.

Before the Pioneer 10 and 11 passages there were serious concerns that spacecraft might not be able to pass safely through the main asteroid belt. Much has been learned about the density of material in the belt since the space age began. Thus far, no spacecraft sent into the belt has experienced serious damage from an impact with asteroidal material or an actual asteroid body. Minor hits on dust detectors have been recorded, however. Robotic spacecraft investigations have provided much information about the nature of the various types of asteroids, as has analysis of meteorites found on Earth that are believed to have come from certain asteroids.

The NEAR spacecraft was launched on February 17, 1996, and was directed toward a rendezvous with the asteroid Eros three years later. Eighteen months out from Earth, NEAR flew by the asteroid Mathilde. It successfully reached Eros, and for well over a year NEAR orbited Eros at varying altitudes providing high-resolution images of the surface of this S-type asteroid. After completing its primary mission, NEAR gently touched down on Eros on February 12, 2001. A total of 69 high-resolution images of the asteroid's surface were taken on the way down during a soft-contact landing. The final picture was taken at an altitude of 130 meters and covered an area of 6 meters by 6 meters. Within that final frame was a portion of a 4-meter-wide boulder, as well as evidence of a dusty surface pocked with small rocks and tiny craters. Much to the surprise of the Johns Hopkins University Applied Physics Laboratory research team controlling the spacecraft, NEAR survived its landing and transmitted data back to Earth for two weeks before falling silent. The team was lucky in that the spacecraft's antenna pointed toward Earth and the solar arrays faced partially toward the Sun after impact.

The next step in spacecraft-based investigations of asteroids is the Dawn mission, a robotic probe designed to orbit two different bodies. Dawn's mission is to visit the two largest asteroids, Ceres and Vesta. By comparison with Ceres, the asteroid upon which NEAR settled was a tiny speck. Ceres is a spherical body with a diameter of 960 kilometers. Indeed under an official review of classification for solar-system objects, Ceres is now officially designated a dwarf planet—a characterization it shares (much to the displeasure of many in the scientific community) with Pluto, which was demoted from full planet status to that of a dwarf. To accomplish its mission on a minimum of propellant, Dawn is outfitted with ion propulsion similar to that demonstrated by the Deep Space 1 spacecraft. To achieve its science goals, Dawn is outfitted with a framing camera, a mapping spectrometer, and a gamma-ray and neutron spectrometer. The goal is to image the surface of these two large asteroids and to determine their composition.

Dawn launched on September 27, 2007, and was set up for a gravity assist from Mars in early 2009. Arrival at Vesta was planned for September, 2011. The ion propulsion system would then break Dawn out of Vesta orbit in April, 2012, and send the spacecraft toward a rendezvous with Ceres in February, 2015. Assuming the spacecraft remains healthy and propellant is available when the primary mission ends in July, 2015, Dawn could be redirected to other asteroids within reach.

Samples of rocks believed to come from various portions of the asteroid belt fall on Earth regularly and have been subjected to intense study. The next step in asteroid investigation would be to return pristine samples of asteroids so the asteroid samples are not altered on their outer layers by passage through Earth's atmosphere. A robotic mission to collect and then return samples from any asteroid is possible with contemporary technology.

Perhaps the greatest potential for insight into the nature of asteroids would be a human expedition to such a body. Shortly after the adoption of the Vision for Space Exploration in 2004, National Aeronautics and Space Administration (NASA) entertained the possibility of sending a crewed Orion Crew Exploration vehicle into deep space for a rendezvous with an asteroid. In terms of propulsion requirements, it is slightly less intensive to send a piloted spacecraft to a near-Earth asteroid than to the Moon. Such a mission could take at least six months to reach a target and up to another year to return to Earth. It could return to Earth large amounts of carefully selected asteroid samples for detailed analysis. The potential for gathering information that might be used someday to divert an asteroid that threatened to impact Earth would be tremendous.

CONTEXT

Asteroids usually cannot be seen with the unaided eye, but they provide important clues for understanding planet formation: They can have major effects on the Earth and, in fact, have had such effects during the planet's history. At one time, it was assumed that the asteroid belt was formed by the breakup of a planet between Mars and Jupiter. However, the combined mass in the belt is much less than that of any planet (only 0.04 percent Earth's mass), and the observed differences in the composition of asteroids at different locations in the belt make it unlikely that they all came from the same planet-sized object. It now appears that asteroids are original debris that was left over after planet formation and that has undergone complex processes such as collisions, fragmentation, and heating. Apparently, strong tidal forces caused by Jupiter's large mass prevented small bodies between it and Mars from combining to form a single planet in their region.

It appears, therefore, that asteroids are among the oldest objects in the solar system, left over from the time immediately before planet formation concluded. Studies of these objects should provide clues to the structure and composition of the primitive solar nebulae. Different types of asteroids found in different regions of the solar system support the theory of planetesimal origin through a sequence of condensation from a nebular disk around the Sun. Asteroids farther from the Sun, beyond the main belt, may have contained more ice; those that formed closer, within the belt, may have been primarily stony or stony-iron materials. Some of these planetesimal precursors of asteroids were probably perturbed during close passes by neighboring planets into elongated Apollo-like orbits that cross Earth's orbit. Other objects on similar orbits may have been comets that remained in the inner solar system long enough to lose their volatile ices by evaporation. Processes of collision and fragmentation among these objects provide direct evidence about the earliest forms of matter.

Special interest in Apollo asteroids arises from their potential for Earth collisions. Objects as small as 100 meters hit Earth about once every two thousand years, and the 30 percent that fall on land can produce craters a kilometer in diameter. Such impacts would devastate much wider areas by their shock waves, and dust thrown into the upper atmosphere could have marked effects on climate. Growing evidence suggests that asteroid collisions in the past might have contributed to major extinctions of species, such as the dinosaurs, and perhaps even caused reversals of Earth's magnetism. Thin layers of iridium, often found in meteorites, have been identified in Earth's crust at layers corresponding to such extinctions. Satellite photography has revealed about one hundred apparent impact craters on Earth with diameters up to 140 kilometers. It is likely that many more succumbed to processes of erosion. Knowledge of Apollo orbits might make it possible to avoid such collisions in the future.

Asteroids also offer the possibility of recovering resources with great economic potential. Some contain great quantities of nickel-iron alloys and other scarce elements; others may yield water, hydrogen, and other materials useful for space-based construction. Estimates of the economic value of a kilometer-sized asteroid reach as high as several trillion dollars. A well-designed approach to space mining might someday help to take pressure off Earth's ecosystem by providing an alternative to dwindling resources, and space-borne manufacturing centers might alleviate pollution on Earth.

Joseph L. Spradley and David G. Fisher

FURTHER READING

Barnes-Svarney, Patricia. *Asteroid: Earth Destroyer or New Frontier?* New York: Basic Books, 2003. In-depth coverage of technical issues about asteroids necessary to understand the danger that an impact on Earth represents. Makes connections to science-fiction stories involving such disasters and allows the reader the ability to determine what is often incorrectly portrayed in doomsday documentaries and fiction with asteroid impact themes.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. A richly illustrated summary of early space-age discoveries that radically revised knowledge of the solar system. Discusses the various types of asteroids.

Bobrowsky, Peter T., and Hans Rickman, eds. *Comet/Asteroid Impacts and Human Society: An Interdisplinary Approach*. New York: Springer, 2007. Suitable for an interdisciplinary college course at the introductory level about the science and societal issues related to an impact on Earth by a near-Earth asteroid or comet.

Gehrels, Tom, ed. *Asteroids*. Tucson: University of Arizona Press, 1979. A classic, authoritative and comprehensive book on asteroids available in English. It contains about fifty articles on every aspect of asteroid research, including extensive references to original research papers. Most articles are technical, but the first seventy-five pages provide a readable introductory survey. Tabulations in the last section provide data of various kinds on all asteroids that have been studied.

Harland, David H. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. New York: Springer Praxis, 2000. Provides virtually all of NASA's press releases and science updates during the first five years of the Galileo mission in a single work, including Galileo's encounters with asteroids. Includes an enormous number of diagrams, tables, lists, and photographs.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Results for the entire NEAR mission are presented. Additional material relating to asteroids is included in chapters on comets, meteorites, planetary evolution, and cratering. An appendix on planetary data includes some asteroid data for comparison, and an extensive bibliography includes about seventy entries on asteroids.

Lewis, John S. *Rain of Iron and Ice: The Very Real Threat of Comet and Asteroid Bombardment*. New York: Basic Books, 1997. A comprehensive survey of meteorites, impact-cratering processes, and the concept of cataclysm. About the latter, the book provides a historical look at the change in scientists' belief in uniformitarianism to their recognition of catastrophism. Plans for preventing a major impact are discussed.

Time-Life Books. *Comets, Asteroids, and Meteorites*. Alexandria, Va.: Author, 1990. Heavily illustrated in color, offers an excellent collection of photographs of comets, pictures of meteorites, and descriptions of asteroids.

See also: Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Impact Cratering; Kuiper Belt; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Planetary Interiors; Planetary Orbits; Planetary Orbits: Couplings and Resonances; Pluto and Charon; Solar System: Origins.

Auroras

Category: Earth

Auroras, commonly called the northern and southern lights, are caused by geomagnetic activity taking place in a planet's atmosphere. By understanding auroras, scientists can gauge the effects of solar activities on planetary environments.

OVERVIEW

Auroral phenomena were first observed on Earth. Only later were such phenomena detected on other planets in the solar system. "Aurora" is a general term for the light produced by charged particles interacting with the upper reaches of the Earth's atmosphere. The term "aurora borealis" specifically refers to the northern dawn, or northern lights; "aurora australis" refers to the southern lights. Auroras appear in an oval girdling the Earth's geomagnetic poles, where magnetic field lines become nearly perpendicular to the surface. In this region, the Earth is not shielded from the space environment as it is at lower latitudes, where magnetic field lines can be almost parallel to the surface. Thus, electrons and ions moving along magnetic field lines can strike the atmosphere directly. Normally, the auroral oval is located about 23° from the north magnetic pole and 18° from the south magnetic pole. Because the north magnetic pole is located in Greenland, the oval is offset toward Canada and away from Europe. Generally, auroras appear at altitudes between 100 and 120 kilometers high, in sheets 1 to 10 kilometers thick and several thousand kilometers long.

The auroral oval is a product of the Earth's magnetic field and is driven by the Sun's output of charged particles. The oval can be enlarged as far north or south as 20° latitude; its normal range is around 55-60°. These variations in range and intensity have been correlated with sunspots, showing that solar activity is the engine that drives auroras and other geomagnetic disturbances. Additionally, scientists usually describe auroral activity in terms of local time relative to the Sun rather than the geographic point over which it occurs. Thus, the Earth can be considered to be rotating beneath auroral events (even though the shape of the oval remains skewed). The first indication that auroral displays might be linked to solar activity came in 1859, when Richard Carrington observed an especially powerful solar flare in white light. A few hours later, he observed a strong auroral display and suspected that the two might be linked.

Electrons impinge upon the upper atmosphere from this field-aligned current, moving in a helical path about Earth's magnetic field lines. The helix of electrons trapped in the Earth's magnetic field will become more pronounced as they approach the poles, until finally their direction is reversed at the "mirror point" and they are reflected back to the opposite pole. Motion back and forth is quite normal. If the electrons are accelerated down into the ionosphere, they encounter oxygen atoms and nitrogen molecules. These collisions will release Bremsstrahlung (braking) radiation. These X rays are absorbed by the atmosphere or radiated into space. The oxygen is dissociated from molecular oxygen and then ionized and electrons freed. Those oxygen ions radiate light when neutralized by free electrons. Nitrogen either is excited and radiates when it returns to the "ground" state or is dissociated and excited. As the atmosphere dwindles gradually into the vacuum of space, starting at about 60 kilometers above the surface of the Earth, the atmosphere forms an electrified layer called the ionosphere, where oxygen and nitrogen molecules are dissociated by sunlight. Because many of these free atoms and molecules are also ionized by sunlight, electric fields and currents move freely, although the net electrical charge is zero.

Auroral structure varies widely. Three major forms have been discerned: quiet or homogeneous arcs, rayed arcs, and diffuse patches. Homogeneous arcs appear as "curtains" or bands across the sky. They sometimes will occur as pairs, and rarely as sets of parallel arcs. They have also been described as resembling ribbons of light. The lower edge of the arc will be sharply defined as it reaches a certain density level in the atmosphere, but the upper edge usually simply fades into space. Pulsating arcs vary in brightness, as energy is pumped in at different rates. Also in the category of quiet arcs are diffuse luminous surfaces, which appear like clouds and have no defined structure; they may also appear as a pulsating surface. Finally, the weakest homogeneous display is a feeble glow, which actually is the upper level of an auroral display just beyond the horizon.

Auroras with rays appear as shafts of light, usually in bundles. A rayed arc is similar to a homogeneous arc but comprises rays rather than evenly distributed light. The formation and dissipation of individual rays may produce the illusion that rays are moving along the length of a curtain. Among rayed arcs, the "drapery" most resembles a curtain and is most active in shape and color changes. If the viewer is directly below the zenith of an auroral event, then it appears as a corona, with parallel rays appearing to radiate from a central point. Drapery displays are often followed by flaming auroras that move toward the zenith.

A view of an aurora, caused by a coronal mass ejection, from the International Space Station. (NASA)

A controversial aspect of auroras is whether they produce any sound. Many observers from antiquity to the present have reported hearing auroras; however, sensitive sound-recording equipment has yet to capture such sounds. This leaves open the question of whether or not the sound is a psychological perception, an electrostatic discharge, or some other phenomenon.

Auroral colors—pink, red, green, and blue-green—are distinct and correspond with specific chemistry rather than being a continuous spectrum typical of thermal radiation. Major emissions come from atomic oxygen (at 557.7- and 630-nanometer wavelengths) and molecular nitrogen (at 391.4, 470, 650, and 680 nanometers). These emissions come from distinct altitudes. The green oxygen line (557.7 nanometers), which peaks at 100 kilometers in altitude, is caused by an excited energy state that relaxes in 0.7 second. The red oxygen line (630 nanometers), which peaks at about 300 kilometers, comes from an excited energy state that relaxes in 200 seconds. While oxygen is energized to this level at lower altitudes, its excited energy will be lost to collisions with other gases long before it can relax naturally. From such comparisons, geophysicists have deduced some of the vertical structure of the atmosphere. X rays and ultraviolet light are also emitted but cannot be detected from the ground. The Dynamics Explorer 1 satellite recorded auroras at 130 nanometers in hundreds of images taken several Earth radii above the North Pole. Numerous Earth-sensing satellites continue to observe and study ionospheric physics and auroral phenomena.

Auroral brightness can vary widely. Four levels of international brightness coefficients (IBCs) are assigned, ranging from IBC I, which is comparable to the brightness of the Milky Way, to IBC IV, which equals the illumination received from a full moon. Auroras usually are eighty times brighter in atomic oxygen than in ionized nitrogen molecules, indicating their origins higher in the atmosphere. Doppler shifting is commonly recorded in the spectra around 656.3 nanometers (hydrogen-alpha), indicating the motion of protons that are neutralized and reionized as they accelerate up or down magnetic field lines. It is theorized that only a small fraction (about 0.5 percent) of the energy that goes into auroras actually produces visible light. The remainder goes into radio waves, ultraviolet rays, and X rays, and into heating the upper atmosphere.

Images from the Dynamics Explorer satellite confirmed the indication by ground-based camera chains that auroras are uneven in density and brightness. One image, for example, shows that auroras thin almost to extinction on the dayside but expand to several hundred kilometers in thickness between about 10:00 P.M. and 2:00 A.M. local time. "Theta" auroras have been recorded in which a straight auroral line crosses the oval in the center, giving the appearance of the Greek letter Θ. This phenomenon may be caused by the splitting of the tail of plasma sheets which extends well into the tail of the magnetosphere, or by the solar wind's magnetic field when it has a direction opposite to that of the Earth.

Photographs taken by spin-scan auroral imagers aboard the Dynamics Explorer 1 satellite show that auroral substorms start at midnight, local time, and expand around the oval. Observations of hundreds of substorms show that they have the same generalized structure but that no two are alike. The satellite imager also showed expansions and contractions in the aurora in response to changes in the interplanetary magnetic field and solar wind. As solar wind plasma meets the Earth's magnetosphere, a shock wave is formed, and the wind is diverted around the Earth. This diversion compresses Earth's magnetic field on the sunward side, while it extends like a comet's tail on the nightside. When the field of the solar wind is oriented toward the south, its field lines reconnect with the field lines of the Earth and allow protons and electrons to enter the magnetosphere; they are normally blocked when the field is oriented to the north.

Auroral activity is strongly driven by the solar wind. If the magnetic field of the wind points north—aligned with the Earth's magnetic field—then the auroral oval is relatively small, and its glow is hard to see. When the solar wind's magnetic field reverses direction, a substorm occurs. The oval starts to brighten within an hour, and bright curtains form within it. At its peak, the oval will be thinned toward the noon side and will be quite thick and active on the midnight side. As the storm subsides and starts to revert to normal about four hours after the field is reversed, the aurora dims and curtains form. Finally, a large, diffuse glow covering the pole may be left as the field becomes stronger in the northward direction.

Two curtain-patterned "dueling auroras," as seen over the Yukon in October, 2001. (©Phil Hoffman/Courtesy, NASA)

The flow of the solar wind past the magnetosphere generates massive electrical currents, which flow mostly from one side of the magnetosphere to the other. Some of the currents, however, connect down the Earth's magnetic field, into and through the auroral oval. Because an electric current is caused by the flow of charged particles, in the process electrons are brought directly into the ionosphere around the poles. Primary currents enter around the morning side and exit around the evening side. Secondary currents flow in the opposite direction. Changes in electrical potential of the magnetosphere, as when it is pumped up by particles arriving in the solar wind, will force the electrons through the mirror point. They are then accelerated deeper into the ionosphere. This auroral potential structure is thin but extends around the auroral oval for thousands of kilometers even to the point of closing in on itself.

Electrojets also form in auroras at low altitudes from an effect known as "E-cross-B drift" (written E × B). At high altitudes, electrons and protons flow freely because there is low gas density and no net current change. At lower altitudes, around 100 kilometers, protons are slowed by collisions with atoms and molecules, but electrons continue to move unopposed. The result is a pair of electrojets, eastward (evening) and westward (morning), which flow toward midnight, then cross the polar cap toward noon. These electrojets heat the ionosphere, especially during active solar periods, when auroras are more intense. This E × B drift in auroral ovals appears to be a major source of plasma for the magnetosphere. It appears that positively charged ions are accelerated upward along the same magnetic field lines, whereas negatively charged electrons precipitate downward. Hydrogen, helium, oxygen, and nitrogen ions compose this flow. Each ion has the same total energy, so their paths vary according to mass. The net effect is that of an ion fountain blowing upward from the auroras which spreads by a wind across the poles.

A little-known subset of the aurora is the subauroral red (SAR) arcs, which appear at the midlatitudes; the magnetic field lines on which they occur are different from those on which auroras appear. SAR arcs, which always emit at 660 nanometers (from oxygen atoms), are dim and uncommon. Modern instrumentation has shown that the SAR arcs are a phenomenon separate from the polar auroras. These arcs may be caused by cold electrons in the plasmasphere interacting with plasma waves or with energetic ions. SAR arcs are believed to originate at an altitude of approximately 19,000-26,000 kilometers during especially strong geomagnetic storms, although the arcs themselves appear at altitudes of around 400 kilometers as the energy from the storm leaks or is forced downward.

Auroras also "appear" in the radio spectrum. Studies in the twentieth century showed that auroras could be sounded by radar at certain frequencies. Satellites in the 1970's started recording bursts of energy in the low end of the AM radio spectrum. This radiation is called auroral kilometric radiation (AKR), because its wavelength is up to 3 kilometers, reflected outward by the ionosphere. Bursts can release 100 million to 1,000 million watts at a time, making them far more powerful than conventional broadcasts by humans. Bursts originate in a region of the sky about 6,400-18,000 kilometers high in the evening sector of the auroral oval. Because the radiation is polarized, it is likely that AKR is caused directly by electrons spiraling along magnetic field lines.

Earth is not the only planet to display auroras. Earthly auroral activity has a power rating of approximately 100 billion watts. Auroral displays on Jupiter were detected by the Voyager 1 spacecraft in 1979. During the 1990's the Hubble Space Telescope performed several investigations of this phenomenon. Auroral activity on Jupiter is hundreds of times stronger than on Earth and also appears always to be energized rather than intermittent, as on Earth. In 2007 the Chandra X-Ray Observatory and Hubble Space Telescope conducted a coordinated study of Jovian auroral activity, seeing the phenomenon in both visible and ultraviolet (Hubble) and X-radiation (Chandra) simultaneously. Observing in multiple wavelengths provides clues to the basic physical process involved that seeing an aurora only in visible light cannot.

Saturnian auroras presented a problem with the current understanding of how auroral displays are produced. In 2005 the Hubble Space Telescope and Cassini spacecraft performed coordinated observations of Saturn's auroral activity. Hubble observed in ultraviolet and visible wavelengths, whereas the Cassini probe in orbit about Saturn recorded radio emissions tied to the auroras. An oddity of Saturn's auroras is that, whereas Jupiter's are not affected very much by the solar wind, Saturn's appears to be. Another is that Saturn's auroras brighten on the portion of the planet where darkness leads to sunlight as the storm increases in intensity, which is not the case for Earthly or Jovian auroras.

Uranus's auroras were detected by the Voyager 2 spacecraft's ultraviolet spectrometer. Contemporary studies of Uranian auroras have been performed by the Hubble Space Telescope. The auroral displays of Uranus mimic those on Earth. Uranus's rings have swept clean much of the region that otherwise would have been a rich collection of trapped charged particles that could be taken down along magnetic field lines into Uranus's magnetosphere to generate auroras. Uranus displays both auroras that are centered about its magnetic poles and the subauroral red arcs seen on Earth. Indeed, on Uranus the SAR arcs are more prevalent than auroras that are centered about the magnetic poles. The latter variety are believed to result from currents that connect Uranus's rather unusual satellite Miranda to the gas giant's magnetic pole. Uranian auroras generate only weak radio signals.

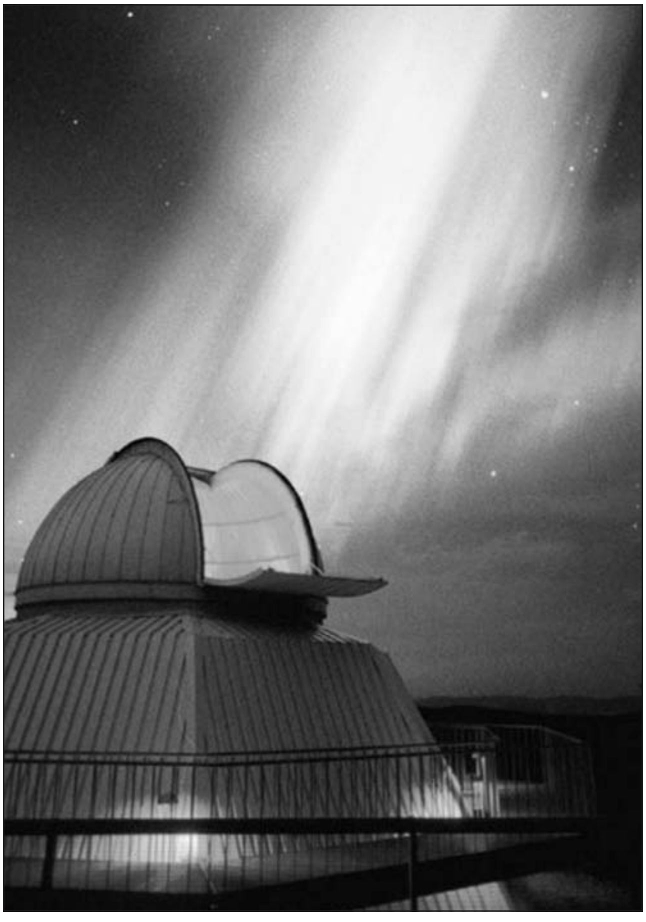
Voyager 2 detected auroral activity on Neptune. Studies of these displays reveal that Neptune's auroral activity is only half as energetic as that on Earth, despite the disparity in size of the two planets. Also, due to the complexity of Neptune's magnetic field, auroral activity is found on Neptune over areas of the planet far from the magnetic poles.

Ironically, even though the planet has no significant magnetic field, Mars also appears to have auroral displays. Data from Mars Global Surveyor and the European spacecraft Mars Express indicated hundreds of aurora-like displays with less dramatic color variations than those observed on other planets in the solar system. Since a planetary magnetic field is not responsible, some researchers suggest that primordial magnetism associated with patches of the planet's crust, particularly in the southern hemisphere, may be involved in this auroral phenomenon. Martian auroral displays show up mostly in the ultraviolet range, with little or no visible counterpart.

METHODS OF STUDY

The space age in some small measure owes its birth to a fascination with auroral phenomena. It was the desire to study and understand the Earth-space interface around the globe at high altitude that resulted in launching the first satellites, during the International Geophysical Year in 1957-1958. Until then, ground-based photography and instrumentation were almost the only methods of studying auroras. Aircraft and rockets played lesser roles. Ground-based instrumentation in the 1940's and 1950's confirmed that auroras were linked to the geomagnetic field, for studies showed that auroras occurred in a circle around the north magnetic pole. Photography of auroral displays has always been difficult because the activity is dynamic, sometimes changing from second to second. Not until the 1950's were electronic devices available to analyze the entire auroral spectrum visible from the ground.

Satellites in the 1970's and 1980's expanded the array of instruments available to investigators. While fields and particles instrumentation has been used to analyze gases and plasmas, imaging instruments have been equally revealing. Notable cameras of various sorts have been carried by Dynamics Explorer 1, the U.S. Air Force HiLat (high latitude) satellite, and the Swedish Viking satellite. In addition, some imaging was performed by polar-orbit weather satellites, but with lower spectral and spatial resolutions. The Skylab crews observed some auroral activity. The space shuttle-based Spacelab 3 crew in 1985 photographed auroras from above the atmosphere. Combining photographs taken a few seconds apart allowed formation of stereo imagery so that the structure could be studied better. In other experiments, small electron guns have been carried into space aboard rockets, on spacecraft, and within the payload bays of space shuttle orbiters. These fire electrons back into the atmosphere in an attempt to generate artificial auroras. An electron gun flew on the first Spacelab (on the space shuttle STS-9 mission) and produced some interesting results. Another such project, on STS-45, was part of the National Aeronautics and Space Administration's (NASA's) INSPIRE program to involve secondary students and undergraduate students in this type of auroral research; unfortunately, although the response of school groups worldwide was strong, the electron gun failed early in the mission.



An aurora, caused by both solar plasma and debris from Comet Swift-Tuttle, appeared over the observatory at Mount Megantic in Quebec, Canada, in August, 2000. (Sébastien Gauthier/Courtesy, NASA)

A key finding from satellite-based research was that auroras are often more active on Earth's dayside, although sunlight and sky completely overwhelm it. Large quantities of radiation are generated in the ultraviolet. This radiation is not seen at the Earth's surface because the atmosphere selectively absorbs such light.

Other planets with magnetospheres also display auroral activity. Such displays are observed almost exclusively by spacecraft either in Earth orbit, like the Hubble Space Telescope, or from spacecraft that either fly by or orbit other planets. The Galileo orbiter routinely picked up auroral activity on Jupiter, and so did the Cassini orbiter on Saturn. Voyager 2 flew past Uranus (1986) and Neptune

(1989) and detected auroral activity on these planets as well.

CONTEXT

Auroras are the most visible manifestation of the interaction between the Earth and space. The study of plasmas has been enhanced by observations made of them. A clear understanding of auroras will provide a means of diagnosing activities in the magnetosphere and the effects of solar activities on the terrestrial environment.

Auroras also serve as a means to study the magnetohydrodynamics of stars and planets. The physics of planetary auroras are essentially the same as that of Earth's auroral activity, although the energies and chemistries involved may be vastly different. Thus, terrestrial auroras can serve as a laboratory for testing basic theories. Planets with magnetic fields also have auroral activity. Much of Jupiter's radio noise is caused by auroral kilometric radiation, and the Einstein Observatory recorded X rays that apparently came from Bremsstrahlung radiation in the Jovian atmosphere. The Voyager 1 spacecraft observed a 29,000-kilometer-long aurora on the nightside of Jupiter, as well as lightning pulses in and above the clouds coincident with the auroral activity. Voyager 2 did too, and it continued on to the rest of the gas giants in the solar system to do the same. Follow-on spacecraft investigations of these planets, and also on Mars, continued to devote considerable effort to characterizing and understanding the nature of auroral activity production. Each planet's auroral activity provides insight into the nature of that planet's magnetic field and also the plasma environment about it.

Dave Dooling

FURTHER READING

Akasofu, Syun-Ichi. "The Dynamic Aurora." *Scientific American* 260 (May, 1989). A detailed, college-level treatment of auroras, written by a physicist who is generally accepted as a world expert.

Bone, Neil. *The Aurora: Sun-Earth Interaction*. New York: John Wiley, 1996. One volume in the Ellis Horwood Library of Space Science and Space Technology. Devoted to describing the electrodynamics of the Sun-Earth environment that produce auroral displays.

Bothmer, Volker, and Ioannis A. Daglis. *Space Weather: Physics and Effects*. New York: Springer Praxis, 2006. A selection from Springer Praxis's excellent Environmental Sciences series, this is an overview of the Sun-Earth relationship and provides a historical and technological survey of the subject. Projects the future of space weather research through 2015 and includes information about contemporary spacecraft.

Delobeau, Francis. *The Environment of the Earth*. New York: Springer, 1971. A technical description of the terrestrial environment, written as a reference for space scientists. Although the work is dated by subsequent discoveries, its description of auroral chemistry is still valid.

Dooling, Dave. "Satellite Data Alters View on Earth-Space Environment." *Spaceflight* 29, suppl. no. 1 (July, 1987): 21-29. An article focusing on the exploration of the magnetosphere by the Dynamics Explorer satellites, with details on auroral imaging and radiation.

Moldwin, Mark. *An Introduction to Space Weather*. Cambridge, England: Cambridge University Press, 2008. This text introduces space weather, the influence the Sun has on Earth's space environment, to the nonscientist. Discusses both the scientific aspects of space weather and issues of technological and societal import.

Savage, Candace. *Aurora: The Mysterious Northern Lights*. New York: Firefly Books, 2001. Provides a history of scientific investigation of auroral phenomena. Heavily illustrated with auroral displays.

See also: Coronal Holes and Coronal Mass Ejections; Earth-Moon Relations; Earth-Sun Relations; Earth's Atmosphere; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Eclipses; Greenhouse Effect; Interplanetary Environment; Neptune's Magnetic Field; Planetary Magnetospheres; Saturn's Magnetic Field; Solar Flares; Solar Wind; Uranus's Magnetic Field; Van Allen Radiation Belts; Venus's Atmosphere.

B

Big Bang

Category: The Cosmological Context

The big bang theory was developed to explain the origin of the expanding universe, uniting cosmology with general relativity and elementary particle physics. About 13 to 14 billion years ago, an explosion called the big bang created energy and matter, space and time. Ever since, space has been expanding with time, carrying matter and electromagnetic radiation with it. As space has expanded, its contents have evolved.

OVERVIEW

Sir Isaac Newton's law of universal gravitation led him to suggest that a static universe with a finite distribution of stars would collapse, but that an infinite universe could be stable. The possibility of an expanding universe, however, is contained within Albert Einstein's general theory of relativity, which he published in 1915. In 1917, Einstein himself found that his general theory in its original form would not permit a static universe. Because the scientific consensus then was that the universe on the large scale is static and unchanging, Einstein added an arbitrary constant (later called the cosmological constant) to his field equations to allow static solutions. Physically, the cosmological constant represents a long-distance repulsion that would balance gravitational attraction on a cosmic scale and thus permit a static universe.

Just five years later, in 1922, the Russian mathematical physicist Alexander Alexandrovich Friedmann found two solutions to the original general relativistic field equations (without the cosmological constant) in which the universe initially expands with time. In one (called "open"), the universe continues to expand forever. In the other (called "closed"), the universe expands to some maximum size, after which it contracts.

In 1927, the Belgian priest and cosmologist Abbé Georges Lemaître independently derived the same two solutions to the field equations of general relativity that Friedmann had obtained earlier. However, Lemaître went further, speculating about the origin of the expansion. Extrapolating backward in time, he realized that everything in the universe would come together at the same time in the distant past, thus pointing to a unique beginning of the universe. He envisioned all matter and space compressed into a "primeval atom" that split into all the atoms of all the elements present in the universe. An enormous explosion initiated the expansion of space and its fragmented matter. As he described the aftermath,

> The evolution of the world could be compared to a display of fireworks just ended—some few red wisps, ashes, and smoke. Standing on a well-cooled cinder, . . . we try to recall the vanished brilliance of the origin of the worlds.

Today we know that the chemical elements could not have been created the way Lemaître proposed. However, Lemaître's basic idea was prophetic. Many years later, the explosive origin of the universe was dubbed the "big bang." Just before his death in 1966, Lemaître learned of the discovery of the cosmic microwave background, which is greatly redshifted radiation emitted just a few hundred thousand years after the big bang—the "vanished brilliance of the origin of the worlds" about which he had speculated so many years earlier.

Observational confirmation that the universe actually is expanding came in 1929, when Edwin Hubble, assisted by Milton Humason, showed there is a correlation between galaxy distances and the redshifts of their spectra; the farther away a galaxy is, the more its spectrum is redshifted. The cause of this redshift, termed cosmological, is the expansion of the universe. As the universe expands, wavelengths of electromagnetic radiation are stretched by the expansion, so visible light is shifted toward longer, redder wavelengths. (The term "redshift" has come to be applied to a shift to longer wavelengths of any part of the electromagnetic spectrum.)

Starting in 1935, Friedmann's student George Gamow began work on more rigorously developing Lemaître's hypothesis of an explosive origin. Gamow proposed that the very dense initial state would have been very hot, and the universe cooled as it expanded. In 1946, he suggested that the primordial substance, which he called "ylem," had consisted of neutrons at a temperature of about 10 billion degrees, some of which decayed during the early stages of expansion to form protons and electrons. Successive interactions of the neutrons and protons then led to the formation of all the chemical elements by nuclear fusion reactions while the early universe still was very hot and dense. Gamow worked out the details of this nucleosynthesis of all the chemical elements with his colleague Ralph A. Alpher at George Washington University. Before they published their results in 1948, Gamow persuaded Hans Albrecht Bethe, the physicist who first described nuclear fusion reactions in stars, to allow them to add his name to their paper to make the list of authors "Alpher, Bethe, Gamow," a wordplay on the first three letters of the Greek alphabet. This came to be referred to as the alpha-beta-gamma theory of the origin of the universe and its chemical elements. (Today we know that the early universe cooled too quickly

George Gamow: Physicist, Cosmologist, Geneticist

Born March 4, 1904, in Odessa, Russia, George Gamow started his scientific career as a boy, when his father gave him a telescope for his thirteenth birthday. Little did his father know that his son would one day become one of the greatest scientists of the twentieth century.

(AP/Wide World Photos)

After graduating from the University of Leningrad in 1926, Gamow went to Göttingen, a center for the study of the new quantum mechanics. At this time, natural radioactivity was the focus of research of many of the great physicists of the day, from the Curies to Lord Rutherford, and Gamow was particularly interested in its relationship to the atomic nucleus. In 1928, he made his first great contribution when he described quantum tunneling of alpha particles to explain the radioactive process of alpha decay. His investigation of the atomic nucleus would take him to Copenhagen, where he worked under Niels Bohr laying the theoretical groundwork for nuclear fusion and fission.

During the 1930's, Gamow taught at universities in Copenhagen, Leningrad, Cambridge, Paris, and the United States. In Washington, D.C., he and Edward Teller worked on the theory of beta decay. He also turned his attention to astrophysics and the origin of the elements. This work led to his 1948 proposal of the "big bang" theory of the universe, for which he is best known.

Gamow was more than a theoretical physicist, however: Known for his sense of humor and revered by his students, he was also devoted to education. His "Mr. Tompkins" series used science fiction to explain difficult science in a way that anyone—including Tompkins, whose attention span was notoriously short—could understand. In 1954, inspired by the Watson-Crick DNA model, he theorized that the order of the DNA molecules determined protein structure. The problem, as he saw it, was to determine how the four-letter "alphabet" of nucleic acid bases could be formed into "words." His "diamond code" paved the way for Marshall W. Nirenberg to crack the genetic code in 1961.

In 1956, Gamow settled in Boulder to teach at the University of Colorado. That year, he received UNESCO's Kalinga Prize for his popularization of science, and two years later he was married (a second time) to Barbara "Perky" Perkins, who initiated the George Gamow Lecture Series after his death, in 1968.

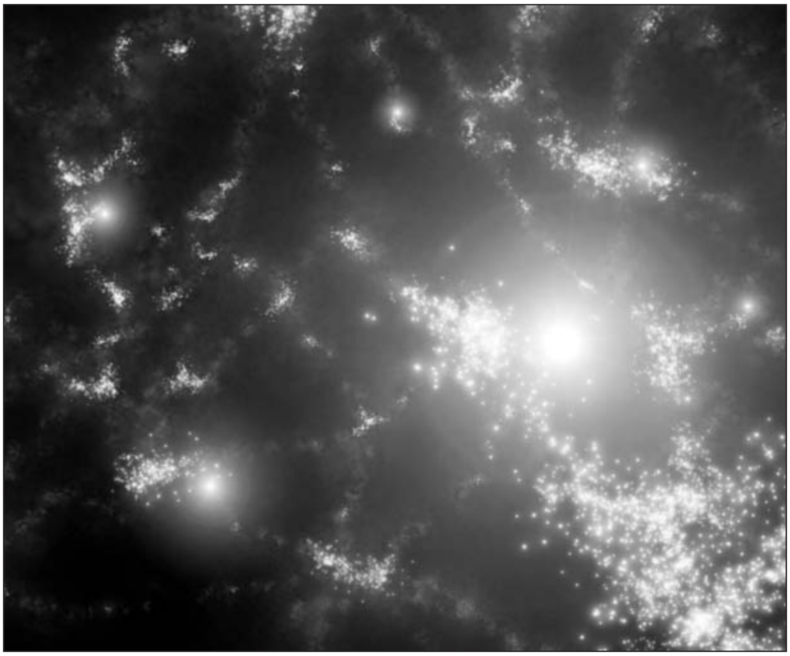
for most of the chemical elements to have formed then; almost all the atoms heavier than helium were formed later by nuclear fusion reactions in stars.)

Gamow and his associates tried to work out other physical processes that would have occurred in the intensely hot, compressed fireball from which the universe expanded. In the same year 1948, Alpher and Robert C. Herman (another of Gamow's colleagues) published a further analysis that predicted a cosmic background radiation left over as a kind of relic from the early hot, dense universe. Because of the expansion of the universe and the corresponding redshift of this radiation, they predicted that it would have cooled from an initial high temperature to only about 5 kelvins at the present time. Since such radiation would be in the microwave part of the spectrum, they had no way of detecting it then, and their prediction was forgotten until the 1960's.

In the 1960's, a team of physicists at Princeton University—Robert H. Dicke, P. J. E. Peebles, P. G. Roll, and David T. Wilkinson—began planning to build an instrument to detect the cosmic background radiation predicted by Alpher and Herman. However, it was accidentally discovered first by Arno A. Penzias and Robert W. Wilson at Bell Telephone Laboratories in 1965. They were using a large radio horn antenna as part of a communication satellite program when they detected microwave radiation coming uniformly from all directions and corresponding to a temperature of about 3 kelvins. Since the signal was so uniform, they thought it might be due to some instrumental noise. Pigeons roosted in the antenna, and Penzias described "a white sticky dielectric substance coating the inside of the antenna." Chasing away the pigeons and cleaning out their droppings did not get rid of the signal. Eventually it was identified by Dicke and his colleagues at Princeton University as the relic radiation from the primeval fireball predicted by Alpher and Herman. In 1978, Penzias and Wilson received the Nobel Prize in Physics for their serendipitous discovery that provided convincing confirmation of a big bang origin for the universe.

Measurements from Earth-orbiting spacecraft such as the Cosmic Background Explorer (COBE), launched in 1989, and the Wilkinson Microwave Anisotropy Probe (WMAP), launched in 2003, as well as high-altitude balloons launched from Antarctica, including the Balloon Observations of Millimetric Extragalactic Radiation and Geomagnetics (BOOMERANG) project, have shown that the microwave background spectrum perfectly fits a thermal radiator at a temperature of 2.73 kelvins. The shape of the spectrum is exactly what would be expected for radiation from the early universe, when matter and radiation were in thermal equilibrium; the shape of the original blackbody spectrum has been preserved during the subsequent expansion and cooling. However, the background radiation is not precisely uniform in all directions. There are temperature variations of up to about 0.00001 kelvin over regions with an angular size of about one degree of arc. These variations are thought to represent slight differences in density in the early universe that ultimately produced the "lumpy" universe of clusters of galaxies that we observe today.

The equations that describe the expansion of the universe can be extrapolated back to very early times of incredibly high densities and temperatures, but only to about 10^{-43} second (called the Planck time). Before that time, conditions were so extreme (for example, temperatures in excess of 10^{32} kelvins) that the current understanding of physics breaks down. Scientists believe that, at that time of high temperatures and correspondingly high energies, the four fundamental forces of nature—gravity, strong nuclear, weak nuclear, and electromagnetic—were indistinguishable from one another, or "unified." However, physicists have no workable "theory of everything" (TOE) to describe this unification of forces. At about the Planck time, gravity would have separated (or "frozen out") from the other three forces as temperature and energy decreased. After about 10^{-35} second, when the temperature had decreased below about 10^{28} kelvins, the strong nuclear force separated from the other two. The energy released by this "freeze out" of the strong nuclear force may have initiated a brief period of cosmic inflation, during which the universe increased in size by a factor of 10^{50} in 10^{-32} second.



This artist's rendering shows how the universe might have looked shortly after the big bang, when matter began to form into stars. (NASA/JPL-Caltech/R. Hurt, SSC)

High-energy elementary particle physics has been employed to work out more details of the early development of the universe and its contents. The first particles and antiparticles that would have materialized from energy according to ideas about mass-energy equivalence (as expressed by Einstein's famous equation $E = mc^2$) and particle-antiparticle pair creation are not well understood, although unified field theories are beginning to suggest their possible properties. At the very high temperatures of the early universe, particles moved so fast that they avoided any interaction, but as the universe cooled they could interact to produce new forms of matter, leading to an era dominated by quarks. The known laws of physics can account for the particles that would have existed after about 10^{-12} second at a temperature of about 10^{16} kelvins. At that time, space was filled with photons, quarks, and leptons (electrons, neutrinos, and the like), along with their antiparticles.

After 10^{-6} seconds, the universe had cooled enough so that no more quark-antiquark pairs could be created. From then on, quarks and antiquarks mutually annihilated each other, producing a brilliant fireball of gamma-ray photons. Equal numbers of quarks and antiquarks had been produced, but some asymmetry resulted in a slight excess of quarks over antiquarks by about one part in a billion. It seems that the asymmetry may occur in the weak nuclear force, which provides a way for antiquarks to decay but no equivalent way for quarks. After the quark-antiquark annihilations were over, all the antiquarks and most of the quarks were gone, but about one quark in a billion had survived; they combined to form protons and neutrons, which went on to become the matter in the universe today.

Neutrons decayed into protons by emitting electrons and antineutrinos, and protons combined with electrons to form neutrons and neutrinos. They were kept nearly equal in number by thermal equilibrium as long as electrons were abundant. When the universe was a few seconds old and the temperature fell to about 6 billion kelvins, photons no longer had enough energy to produce electron-positron pairs. Soon, all positrons and all but one electron out of a billion had mutually annihilated in another burst of gamma-ray photons. With so few electrons remaining, no new neutrons were formed, and the number of neutrons declined as they decayed into protons.

Before all the neutrons decayed, some joined with protons to form nuclei of deuterium (also called heavy hydrogen). However, while the universe was hot enough, gamma-ray photons could break deuterium nuclei apart. After about three to four minutes, when the temperature had dropped below about 1 billion kelvins, photons no longer had enough energy to break up deuterium nuclei, so they could survive. In rapid succession, the deuterium nuclei then collided with protons and neutrons to form helium nuclei, and soon almost all the remaining neutrons combined to form helium. When this nucleosynthesis began, there was about one neutron for every six protons. Using almost all the neutrons to form ordinary helium nuclei resulted in one helium nucleus (two protons and two neutrons) for every ten hydrogen nuclei (each just a single proton), and this very closely matches the cosmic abundance of helium and hydrogen observed today. A few lithium nuclei and even fewer beryllium nuclei also formed, but the temperature dropped too quickly for there to be time to form heavier nuclei. After about fifteen minutes, the temperature had dropped below 400 million kelvins, and nucleosynthesis ended. (The heavier elements eventually formed at much later times through nuclear fusion reactions in stars.) Throughout the early universe, the radiation density exceeded the matter density, but radiation density decreased more rapidly than matter density. After several thousand years, the two densities were equal, and from that time on, matter has been dominant.

When nucleosynthesis ended, one electron remained for each free or bound proton in hydrogen and helium nuclei, but the universe was too hot for electrons to combine with nuclei to form neutral atoms. Free electrons are very effective at scattering photons, so the universe was opaque to electromagnetic radiation. The universe expanded for several hundred thousand years before it was cool enough (about 3,000 kelvins) for electrons to combine with nuclei to form neutral atoms. When this happened, the lack of free electrons made the universe transparent to electromagnetic radiation, and photons were free to travel through space. This decoupling of matter and radiation was the source of the cosmic microwave background radiation. As the universe continued to expand, it stretched the wavelengths and effectively cooled the primeval "relic" radiation until it reached the present temperature of 2.73 kelvins.

APPLICATIONS

The cosmological interpretation of redshifts attributes these spectral shifts to the stretching of wavelengths of electromagnetic radiation as space expands. The greater the redshift, the more space has expanded since the electromagnetic radiation was emitted, and the farther back in time one can observe. The Hubble law expresses basically the same idea; the greater the redshift, the farther away the sources and the greater the travel time of electromagnetic radiation to reach Earth. This ability to look back in time aids our understanding of distant objects with large redshifts. Such objects are seen as they were billions of years ago during early stages in the evolution of the universe.

For example, quasars (quasi-stellar radio sources) are mostly unresolved sources with very small angular sizes having very large redshifts and often rapid and erratic changes in brightness. If their redshifts are cosmological (a few astronomers dispute this assertion), then they are very distant. This conclusion, coupled with their apparent brightness, implies that they are incredibly luminous. Their rapid changes in brightness mean they are relatively small in actual size. All of this taken together suggests that they likely are extremely energetic compact phenomena in young galaxies, possibly supermassive black holes forming at the centers of developing galaxies.

The cosmic microwave background radiation is extremely uniform in all directions, but it does have a small asymmetry, being slightly warmer systematically in one half of the sky compared to the other half by a maximum of about 0.007 kelvin. This is interpreted as due to the motion of our solar system through the background radiation field; the temperature difference of 0.007 kelvin in thermal radiation corresponds to a spectral Doppler shift produced by a speed of 380 kilometers per second. Presumably this speed is a combination of the movements of our solar system in the Milky Way galaxy, the Milky Way galaxy in the Local Group of galaxies, the Local Group relative to the Virgo supercluster of galaxies, and maybe other motions as well.

The measured temperature of the cosmic microwave background radiation makes it possible to calculate the expected cosmic abundances of light elements: about 74 percent (by mass) hydrogen, 26 percent helium, a thousandth of a percent deuterium, and a millionth of a percent lithium. All these match the measured abundances within the observational uncertainties. Since no other source for deuterium production is known, these measurements provide additional confirmation for the standard big bang model.

The uniformity of the cosmic background radiation implies thermal equilibrium throughout the universe, even in regions so far apart that electromagnetic radiation has not had time to travel from one to the other. In the early 1980's, Alan Guth proposed "inflation" as a solution to this "horizon problem." He suggested a very early period of rapid expansion at an exponential rate, when the universe increased in size by a factor of 10^{50} in 10^{-32} second. Before inflation, the universe would have been small enough for electromagnetic radiation to travel between all parts of it.

Uniformity also raises a "galaxy problem." How could galaxies and stars form in such a uniform universe? Fine-scale sky maps of the cosmic background radiation made with data from sensitive detectors on the COBE and WMAP spacecraft and BOOMERANG balloons show small temperature (and hence density) variations in the early universe that could have grown by gravity to develop into clusters and superclusters of galaxies. The density variations probably are due to small random quantum fluctuations in the very early universe.

The discovery of the W and Z particles in 1983 by Carlo Rubbia and Simon van der Meer provided support for the electroweak theory, which predicts the electromagnetic force and the weak nuclear force become "unified" or indistinguishable from each other at temperatures above about 10^{15} kelvins. The quark theory predicts a weakening of the strong nuclear force at even higher temperatures; grand unified theories (GUTs) propose that the strong nuclear force is unified with the electromagnetic and weak nuclear forces at temperatures above about 10^{27} kelvins. Theories of everything (TOEs) go still further and suggest the unification of gravity with the other three forces at temperatures above about 10^{32} kelvins.

The high temperatures required to unify the forces occurred shortly after the big bang, making the very early universe a high-temperature laboratory in which it may be possible to test these theories. Gravity would have decoupled at the Planck time, 10^{-43} second. The decoupling of the strong nuclear force at about 10^{27} kelvins would have occurred after about 10^{-35} second and may have released the energy that drove the sudden inflationary expansion of the universe. The last decoupling of the weak nuclear and electromagnetic forces would have occurred at 10^{-11} second, when the temperature had cooled to 10^{15} kelvins.

CONTEXT

Several competing theories have attempted to avoid the creation implications of the big bang theory, but they have not been able to sustain successful alternatives. One of the first was Einstein's early attempt to obtain a solution for a static universe and his introduction of an arbitrary cosmological constant to balance gravitational attraction. When it was later shown that his field equations of general relativity without the cosmological constant were compatible with the observed expansion of the universe, Einstein is reported by Gamow to have remarked that the cosmological constant was the greatest blunder of his life. It is ironic that the early era of cosmic inflation and the recently discovered acceleration of the expansion both involve repulsive forces similar to Einstein's original cosmological constant.

The most serious attempt to defeat the big bang theory was the steady state theory of the universe, introduced in 1948. About the same time that Gamow and his colleagues were working out the details of a Lemaître-type explosive origin, the British cosmologists Hermann Bondi, Thomas Gold, and Fred Hoyle were developing an alternative—the steady state continuous creation model. It was Hoyle who coined the term "big bang" as a derogatory name for the Lemaître-type primordial explosion. However, the name is short and catchy, and it was quickly adopted by most astronomers and physicists no matter which side (if either) they supported.

The steady state theory did not invoke a moment of creation for the entire universe but assumed instead the continuous creation of new matter throughout space at a rate that keeps the mean density of the universe constant for all times as the universe expands. Continuous creation would occur so gradually that it could not be observed until enough matter had been created to form stars and galaxies. Such a steady state universe would be infinite and eternal. Ironically, it required the religious idea of creation *ex nihilo* (from nothing) to avoid another religious idea of a unique creation event in the remote past (the big bang).

Although the steady state theory provided the main competition for the big bang theory during the 1950's and early 1960's, it did not stand the test of time. Since stars and galaxies would form throughout space from the continuous creation of new matter, young and old galaxies should exist side by side. This is contrary to the evidence that galaxies all formed at about the same time, and the galaxies that seem to be much younger (such as quasars) are observed only at great distances and hence at great times in the past. The steady state theory was virtually abandoned after the 1965 discovery of the cosmic background radiation, the relic radiation predicted from the big bang fireball. Even Hoyle, chief spokesman for the steady state theory, helped work out some details of the standard model of the big bang in 1967.

One other attempt to avoid a finite age for the universe was the idea of an oscillating universe. If the density of matter in the universe were large enough eventually to reverse its expansion by gravitational attraction, the universe would collapse toward a "big crunch." The oscillating universe theory proposed that another big bang might follow each big crunch, giving rise to a series of oscillations between successive big bangs, extending indefinitely into the past and future. However, such speculation was laid to rest by the discovery in the 1990's that the expansion of the universe is accelerating, so no contraction seems possible.

Joseph L. Spradley and Richard R. Erickson

FURTHER READING

Barrow, John D., and Joseph Silk. *The Left Hand of Creation*. New York: Basic Books, 1983. A readable account of the origin and evolution of the expanding universe by two astronomers with a good grasp of cosmology. Discusses many theories and problems associated with the big bang model, and includes a good glossary of astrophysical terms and an index.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well written college-level textbook for introductory astronomy courses. Two chapters provide a thorough description of the big bang and the evolution of the universe afterward.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. One chapter contains a good description of the big bang and its aftermath.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook, thorough and well-written. One chapter contains a good description of the big bang and its aftermath.

Hawking, Stephen W. *A Brief History of Time: From the Big Bang to Black Holes*. New York: Bantam Books, 1988. A very popular and readable account of the development of cosmology and the big bang theory. Includes a helpful glossary and an index.

Jastrow, Robert. *God and the Astronomers*. New York: Warner Books, 1978. A brief but interesting history of the discovery of the expanding universe and development of the big bang theory. Contains many historical photographs of the originators of the theory and supplements on its theological implications.

Lang, Kenneth R., and Owen Gingerich, eds. *A Source Book in Astronomy and Astrophysics, 1900-1975*. Cambridge, Mass.: Harvard University Press, 1979. This volume contains many of the original articles that established the ideas of the expanding universe and the big bang theory, with good introductory sections for each. Contributors include Einstein, Hubble, Friedmann, Lemaître, and Gamow. Some articles are technical, but much can be understood by the general reader.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses.

Divided into short units on specific topics. Several units provide a thorough discussion on the big bang and the evolution of the universe afterward.

Silk, Joseph. *The Big Bang*. Rev. ed. New York: W. H. Freeman, 1989. A good introduction to the standard model of the big bang theory. Includes a good glossary, index, and a thirty-five-page section on mathematical details.

Trefil, James S. *The Moment of Creation: Big Bang Physics from Before the First Millisecond to the Present Universe*. New York: Charles Scribner's Sons, 1983. A good introduction to the big bang theory. Includes a discussion of grand unification and inflationary theories.

Weinberg, Steven. *The First Three Minutes*. New York: Bantam Books, 1977. An excellent introduction to the details of the standard model of the big bang by a leading theoretical physicist and Nobel laureate.

See also: Cosmic Rays; Cosmology; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; General Relativity; Interstellar Clouds and the Interstellar Medium; Milky Way; Novae, Bursters, and X-Ray Sources; Solar System: Element Distribution; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Stellar Evolution; Supernovae; Universe: Evolution; Universe: Expansion; Universe: Structure.

Brown Dwarfs

Category: The Stellar Context

Between the giant planets such as Jupiter, in which no nuclear reactions occur, and the small red dwarf stars, in which nuclear reactions produce energy, objects exist whose mass is almost great enough to have initiated a few nuclear reactions but which mostly just radiate the heat that nearly ignited them. Known as brown dwarfs because of the feeble infrared light they emit, the first of these was unequivocally identified only in 1995.

OVERVIEW

Brown dwarfs are defined as objects with masses intermediate between stars and planets—not massive enough to fuse ordinary hydrogen nuclei (consisting of single protons, H^1) into helium in their cores as "real" stars do at some stage, but massive enough to generate energy briefly by nuclear fusion of deuterium (heavy hydrogen consisting of a proton and a neutron, H^2). Theoretical calculations indicate that the upper mass limit is about 7 to 8 percent ($1/14$ to $1/14$) the Sun's mass, or about 70 to 80 times Jupiter's mass; above this, ordinary hydrogen fusion occurs. The lower mass limit is estimated to be about 1.0 to 1.7 percent ($1/100$ to $1/60$) the Sun's mass, or about 10 to 17 times Jupiter's mass; below this, no nuclear reactions of any sort can occur and the object simply is a large planet, a "super-Jupiter." Some astronomers prefer to reserve the term "star" for objects massive enough to initiate ordinary hydrogen fusion, and these astronomers call brown dwarfs "failed stars" or "almost stars" or "substellar objects."

Because of their low mass, brown dwarfs have low temperatures by stellar standards. Their surface temperature is 2,000 kelvins (degrees above absolute zero) or less. In contrast, the Sun—by no means a very hot star—has a surface temperature of about 6,000 kelvins. They are called brown dwarfs because, due to their low surface temperature, most of their electromagnetic radiation is in the infrared part of the spectrum; at visible wavelengths they glow faintly with a dim, dark red color. Their diameter is about $1/10$ the Sun's diameter, which makes them about the same size as Jupiter. Their surface area and its temperature determine their luminosity, which ranges from about $1/10,000$ down to $1/1,000,000$ (10^{-4} to 10^{-6}) of the Sun's luminosity.

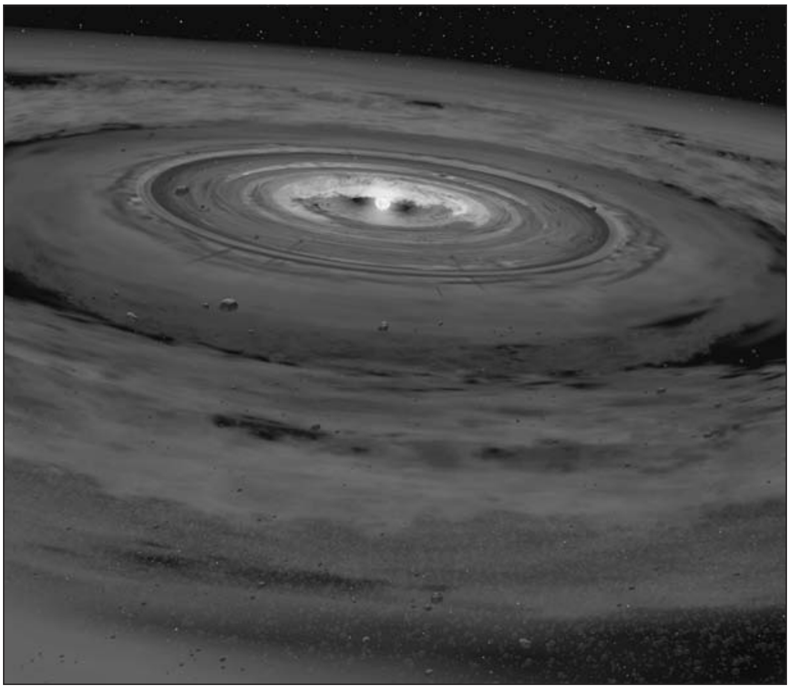
METHODS OF STUDY

Astronomer Jill Tarter coined the name "brown dwarf" in 1975 for hypothetical objects in between stars and planets. In subsequent years, other astronomers predicted their appearance and physical properties, postulating that our Milky Way galaxy contained many of them because slightly more massive red dwarf stars are so abundant.

Tarter's speculation touched off a search for brown dwarfs by many of the world's major observatories. The problems in identifying such objects were formidable. Brown dwarfs are very cool and faint, emitting very weak electromagnetic radiation primarily at near-infrared wavelengths of a few microns (10^{-6} meters). Predicted spectral signatures included absorption bands due to water (H_2O) and methane (CH_4), since such stars would be cool enough for these compounds to form; in both compounds the bonds between the hydrogen atoms and the central oxygen or carbon atom absorb energy in a narrow band of wavelengths within the near-infrared part of the spectrum.

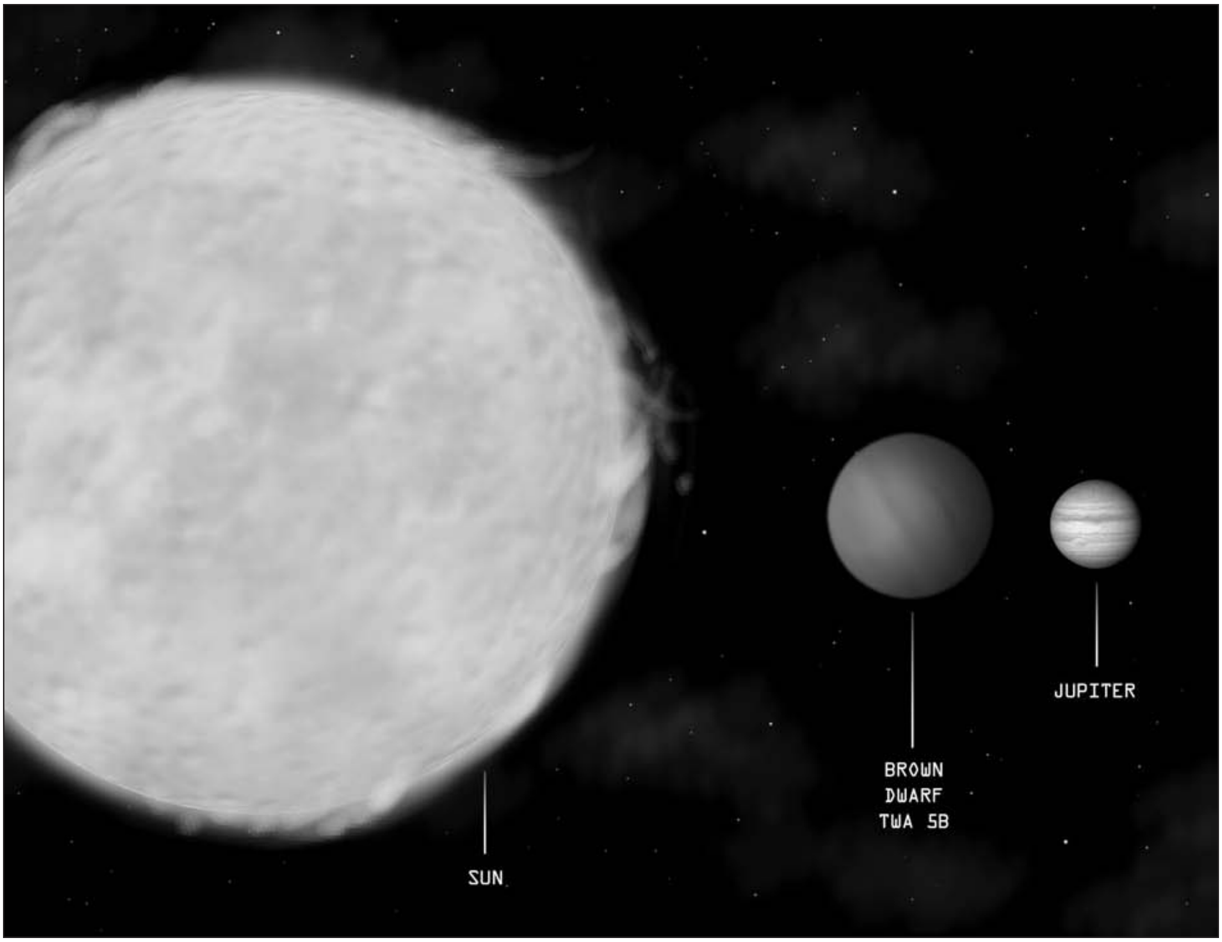
Even more conclusive spectral evidence is provided by the element lithium. Small amounts of the isotope lithium 7 (Li^7) were produced in the big bang. Li^7 can undergo nuclear reaction when bombarded with a proton; its nucleus splits, and two atoms of helium 4 (normal helium, He^4) are formed. This happens, however, only at the temperatures found in "real" stars. A brown dwarf is cool enough for lithium to be consumed only very slowly, if at all. A spectrum showing an absorption feature at the wavelength characteristic of lithium, 0.67 micron, is almost certain confirmation that the object in question is a brown dwarf.

The first bodies to be identified as brown dwarfs were Teide 1 in the Pleiades and Gliese 229B, part of a binary star system with the true star Gliese 229, located in the constellation Lepus (the Hare). Since Gliese 229's distance from the Earth was known, the brown dwarf's distance was also known: about 19 light-years. Teide 1 is much farther away (about 400 light-years) and harder to observe. Although Teide 1 was discussed in the research literature before Gliese 229, it had to await final confirmation while the identification of Gliese 229B became fully established. In 1995 Gliese 229B was the subject of two papers: one published in *Science*, which provided methane spectral evidence, and the other in *Nature*, which provided lithium data. Because the methane absorption is so strong, Gliese 229B is considered to be surrounded by a thick methane atmosphere, somewhat like Jupiter. The lithium absorption was the final piece of evidence, confirming that Gliese 229B is a brown dwarf.



An artist's conception of the brown dwarf OTS 44, about fifteen times the size of Jupiter, at the center of a protoplanetary disk. (NASA/JPL-Caltech)

Enough brown dwarfs have now been discovered that they have been assigned spectral types L and T as an extension of the existing sequence of stellar spectral types O, B, A, F, G, K, and M. Type M refers to the "real" stars with the coolest surfaces, down to about 2,000 kelvins. Type L is applied to brown dwarfs with surface temperatures between 2,000 and 1,300 kelvins. Their spectra are characterized by absorption bands and lines due to water, carbon monoxide, metal hydrides, sodium, potassium, cesium, and rubidium. Type T refers to brown dwarfs with temperatures from 1,300 kelvins on down (perhaps to about 700 kelvins). Their spec-



The brown dwarf TWA 5B (center) is shown with the Sun on the left and Jupiter on the right for size comparison. (NASA/CXC/M. Weiss)

tra are characterized by absorption bands of water and methane. Teide 1 is an example of type L, and Gliese 229B is an example of type T. Our understanding of the cooling rate of brown dwarfs is that they start out as type L, and after no more than about 1 billion to 2 billion years they have cooled down to type T.

CONTEXT

The actual detection of brown dwarfs after their existence was predicted helped fill in theories about the formation of stars and planets. Red dwarf stars (also called red main sequence stars, of spectral type M and luminosity class V) with masses down to about 7 or 8 percent of the Sun's mass are the least massive "real" stars, slowly fusing hydrogen into helium in their cores. Many such stars are known; they are the most common type of star found in our solar neighborhood, and presumably throughout the Galaxy. Then there are the giant planets like Jupiter and Saturn, with masses less than 0.1 percent of the Sun's mass. They have nearly the same elemental abundance as young stars (mostly hydrogen, most of the rest helium, and small amounts of other elements), but they lack the mass to have generated high enough temperatures by gravitational contraction to have initiated hydrogen fusion. Theory suggested that between red dwarf stars and giant planets like Jupiter and Saturn there should exist a class of intermediate objects—objects that generated substantial heat as they first contracted, perhaps enough to fuse deuterium (heavy hydrogen), but not enough to fuse protons (ordinary hydrogen nuclei) into helium. The successful identification of brown dwarfs and the continuing discovery of increasing numbers of them suggest a continuity in the mass distribution function (the number of objects as a function of mass) from stars down to planets. In fact, brown dwarfs may outnumber ordinary stars in the Galaxy.

Robert M. Hawthorne, Jr.

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. College-level textbook for introductory astronomy courses. Has more thorough discussion of brown dwarfs than most.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Several pages refer to brown dwarfs.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook. Thorough and well-written with several pages discussing brown dwarfs.

Marley, M. S., et al. "Atmospheric, Evolutionary, and Spectral Models of the Brown Dwarf Gliese 229 B." *Science* 272 (June 28, 1996): 1919-1921. A fairly technical article, but it provides the actual data upon which the identification of Gliese 229B is based.

Mayor, Michel, and Disier Queloz. "Many Brown Dwarfs Being Found by Lithium Signature." *Sky and Telescope* 93 (February, 1997): 17-23. More discoveries, as in the preceding citation.

_______. "Swiss Find Ten New Brown Dwarfs." *Astronomy* 25 (February, 1997): 24-29. Part of the rush of discovery that followed Teide-1 and Gliese 229B.

Nakajima, T., et al. "Discovery of a Cool Brown Dwarf." *Nature* 378 (November 30, 1995): 463-465. Data on such aspects of Gliese 229B as mass, temperature, and luminosity and a demonstration that this brown dwarf is as far from Earth as its parent star. A technical article.

Reid, Neil, and Suzanne Hawley. *New Light on Dark Stars: Red Dwarfs, Low-Mass Stars, Brown Stars*. 2d ed. New York: Springer Praxis, 2005. A technical description of stars and other objects which are not very luminous. Discusses recent discoveries of brown dwarfs and extrasolar planets.

Rosenthal, Edward D., Mark A. Gurwell, and Paul T. P. Ho. "Efficient Detection of Brown Dwarfs Using Methane-Band Imaging." *Nature* 384 (November 21, 1996): 243-244. Methane data following on Nakajima's article.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses. Divided into lots of short sections on specific topics. Has several pages referring to brown dwarfs.

Tyson, Neil de Grasse. "When a Star Is Not Born." *Natural History* 105 (March, 1996): 62-63. A popular discussion of brown dwarfs and their formation, this article touches on the theoretical question of whether brown dwarfs fit into existing star-formation mechanisms.

See also: Comets; Extrasolar Planets; Extrasolar Planets: Detection Methods; Gamma-Ray Bursters; Infrared Astronomy; Main Sequence Stars; Novae, Bursters, and X-Ray Sources; Nuclear Synthesis in Stars; Protostars; Pulsars; Red Dwarf Stars; Red Giant Stars; Stellar Evolution; Supernovae; Thermonuclear Reactions in Stars; White and Black Dwarfs.

C

Callisto

Categories: The Jovian System; Natural Planetary Satellites

Study of Callisto, Jupiter's outermost natural satellite, has led to insights into the formation of the solar system, the possibilities for extraterrestrial life, and the protection from comet impacts that Jupiter gives to the inner planets of the solar system.

OVERVIEW

Callisto is the outermost of the four major satellites of the "gas giant" planet Jupiter. It was discovered with one of the earliest telescopes by Galileo Galilei in 1610. Hence, it is often referred to as one of the Galilean satellites. Callisto is one of the largest satellites in the solar system, ranking third behind Jupiter's Ganymede and Saturn's Titan. With a diameter of 4,800 kilometers (2,985 miles), it is nearly the size of the planet Mercury. Callisto is also tidally locked to Jupiter, meaning that its "day" is the same length as its month, 16.82 Earth days. As a result, the same side of the satellite always faces Jupiter, just as the Moon always presents the same face toward Earth.

If the Galilean satellites had personalities, Callisto would be a frail old man. Unlike the young and vibrant Io, Callisto has neither volcanoes nor large mountains anywhere on its surface. In fact, its total lack of geological activity, both above and below the surface, means that its surface most likely resembles what the satellite looked like during its formation. This is at least partly due to the lack of tidal forces from nearby Jupiter. The lack of squeezing and pulling from Jupiter's gravity reduced the heat and energy within the satellite, leading to a relatively tranquil geology. This unique surface gives astronomers and geologists a glimpse of not only the primordial Jovian system but also the primordial solar system.

Callisto's surface is twice as bright as Earth's moon but still much darker than the surfaces of its Jovian siblings. The first few kilometers of the surface layer is primarily ice, with a darker material having leaked in at some point. Callisto's surface is uniformly covered in craters and is thought to be the most cratered satellite in the solar system. These impacts are the primary force that has shaped the planet, and sometimes great rings appear around the impact craters. The two largest features, Valhalla and Asgard, are respectively 3,000 kilometers (1,865 miles) and almost 1,600 kilometers (1,000 miles) in diameter. While impacts have been the primary force in shaping Callisto's surface, data from the Galileo space probe in the late 1990's showed that some minor erosion has occurred. This erosion is thought to be carbon dioxide sublimating through cracks in the surface ice.

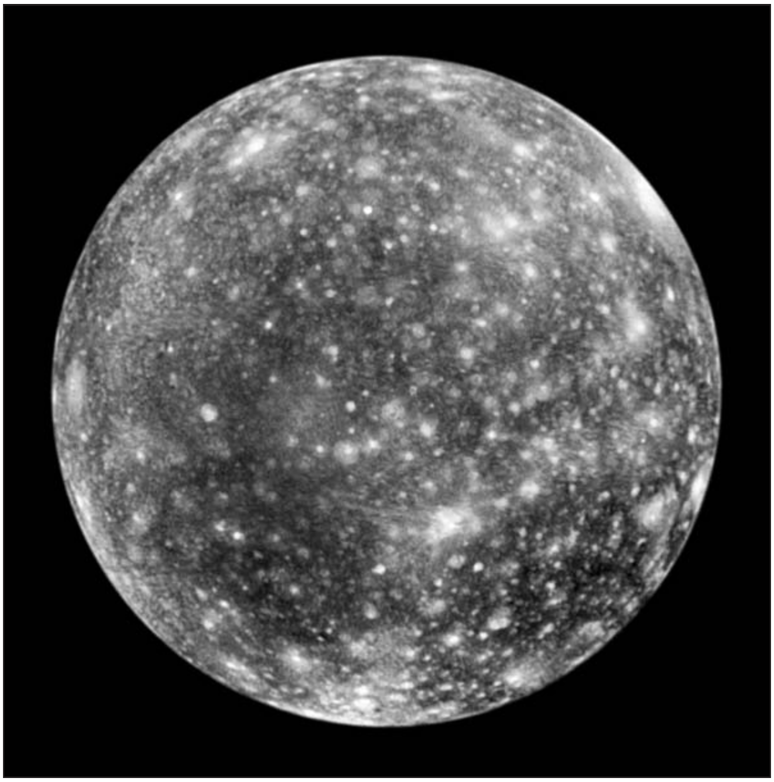
Along with these large impact craters, there are numerous crater chains, or catenae. After the 1979 Voyager flybys, the catenae were thought to be the result of debris from asteroid impacts. This idea was called into question after the spectacular impact of Comet Shoemaker-Levy 9 into Jupiter's atmosphere during late May, 1994. This comet had come within a special distance from Jupiter, known as the Roche limit, and been broken up by the force of gravity. What was once one large comet was now a series of fragments traveling in formation. This event gave credibility to the idea of comets colliding with planets and satellites and has helped to explain Callisto's pockmarked surface.

While the surface has given scientists relatively overt information about the satellite's past, Callisto's interior remains shrowded in mystery and conjecture. With a density of 1.86 grams/centimeter^3, Callisto's density is the smallest of the major Jovian satellites. Scientists at the National Aeronautics and Space Administration (NASA) believe that Callisto is made up of roughly equal parts rock and ice, but the exact internal structure is unclear. Early observations led Galileo scientists to believe that Callisto is undifferentiated, meaning it has the same composition throughout.

Most rocky bodies in the solar system, such as Earth, have multiple layers that form during their creation. Molten materials tend to separate out, or differentiate, due to density. Within Earth, for instance, there is a dense core of iron and some nickel. Moving away from the core are different layers of decreasing density. Initial readings from Galileo showed that this process had not taken place in Callisto. Newer data, from subsequent flybys, do not directly contradict this hypothesis but have made planetary scientists less certain. Further evidence for an undifferentiated interior comes from data showing that Callisto also lacks its own magnetic field, suggesting a lack of a metallic core.

Curiously enough, however, Callisto does alter Jupiter's magnetic field within its vicinity. Because this perturbation in the field arises from increased conductivity within the planet, scientists speculate that a subsurface ocean may exist. Only an ocean with the salinity similar to Earth's oceans could explain the readings.

Callisto also has an extremely thin atmosphere composed primarily of carbon dioxide. With a pressure millions of times lower than Earth's, the atmosphere appeared, based on data from the Galileo flybys of 1998-1999, to have formed relatively recently. These data led scientists to believe that the atmosphere was no more than four years old and due to a combination of processes known as photoionization and magnetospheric sweeping. Photoionization takes place when ultraviolet rays (the same rays that cause sunburns) come in contact with individual carbon dioxide (CO_2) molecules; each CO_2 molecule ejects an electron, similar to the way a solar calculator generates current. Removal of an electron causes the molecule to become charged. Since charges interact with magnetic fields, Jupiter's enormous magnetic field acts like a giant broom and sweeps these ionized particles away from Callisto. Left unchecked, this process would eventually cause Callisto's atmosphere to fade away.

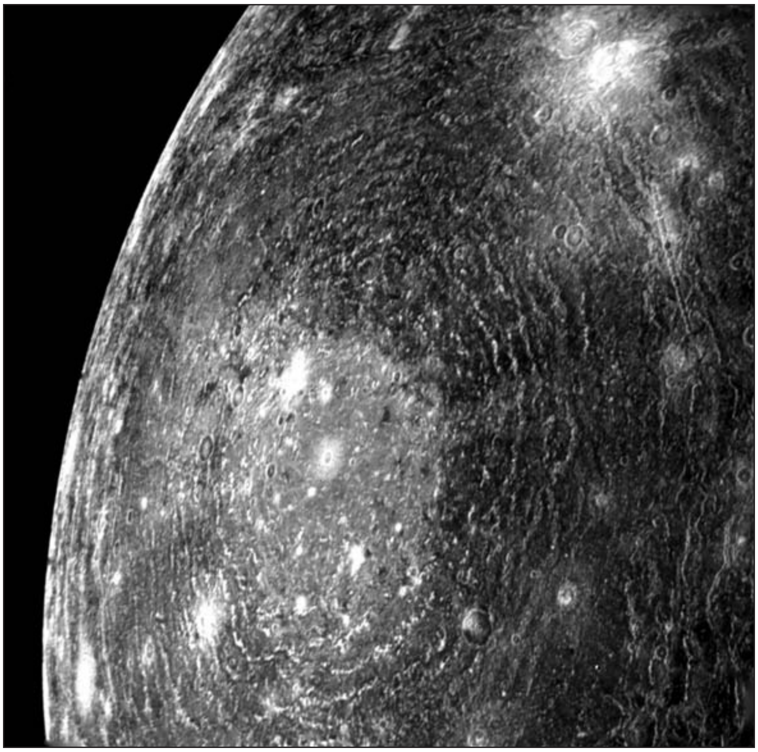


Jupiter's pockmarked moon Callisto, as imaged from the Galileo spacecraft in 2001. (NASA/JPL/DLR)

If the atmosphere is not transient, the carbon dioxide gas must be replenished on a continual basis. The obvious source of CO_2 gas is Callisto's icy surface. This ice would have to be located in a region that is permanently shadowed, away from direct light and protected from ionization. It has also been suggested that much of the carbon dioxide that exists on the satellite's surface, as well as this tenuous atmosphere, comes from the comet impacts that Callisto has sustained.

KNOWLEDGE GAINED

The vast majority of Callisto data comes from the Voyager flybys of the late 1970's and the multiple flybys of the Galileo spacecraft during the late 1990's. Before that, the satellite was, at best, a foggy image in



Taken in 1979 by one of the Voyager spacecraft from about 200,000 kilometers, this image of Callisto shows a multiple-ring basin. (NASA/JPL)

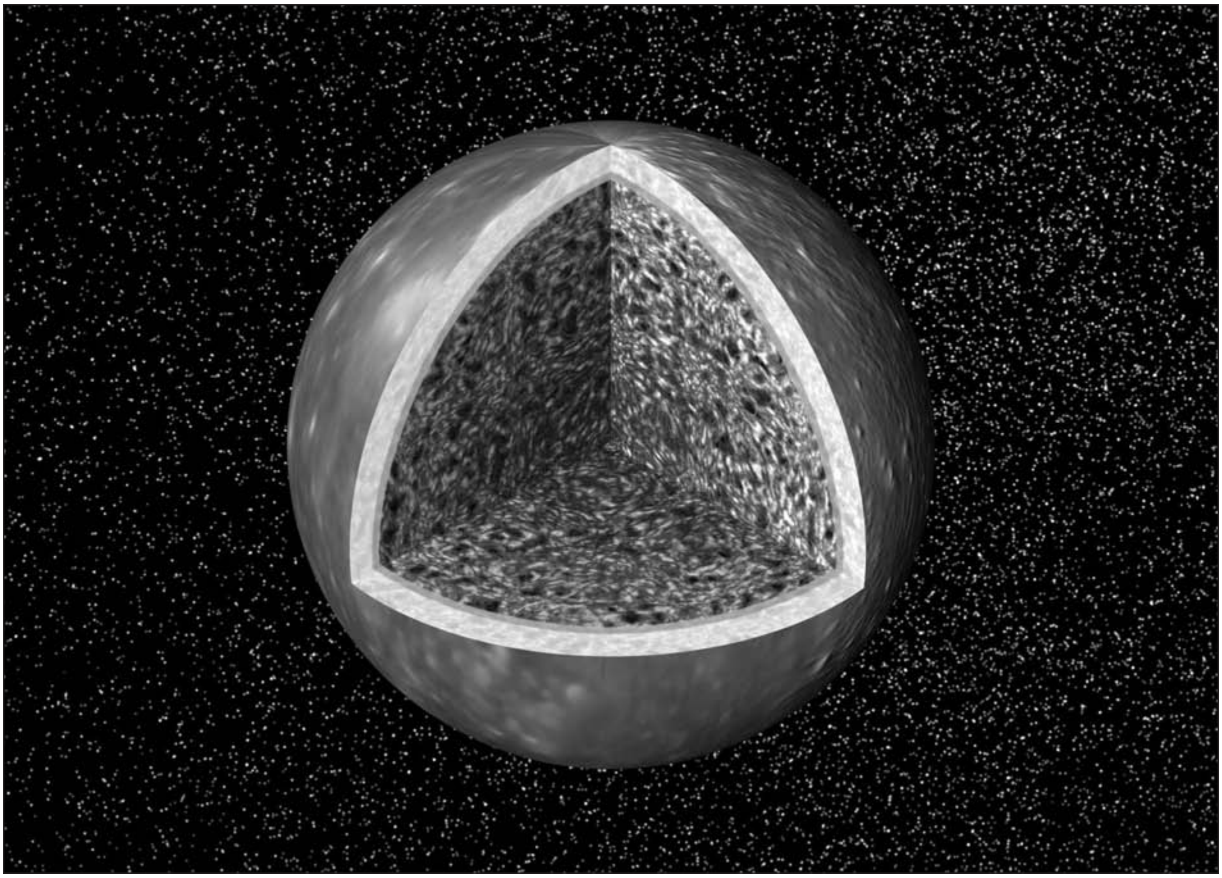
ground-based professional telescopes and a minuscule, but predictable, pinprick of light in backyard telescopes. Even Hubble Space Telescope images taken in October of 1995 showed a blurry surface. Only uncrewed space probes would produce the information needed to gain further understanding.

Both Voyagers 1 and 2, which took images on their way to the outer solar system, revealed a relatively dead world, battered by impact craters. Two decades later, Galileo returned to focus purely on the Jovian system. Its more sophisticated instruments offered higher-resolution imagery, magnetometric information, and spectroscopic information.

Galileo's most significant discovery about Callisto was the possibility of an underground ocean, similar to Earth's oceans. The discovery of water in the solar system is always a major event because it is thought to be an essential ingredient for life. Water was already thought to exist on nearby Europa, and great efforts were made to ensure that Galileo would not contaminate the surface. This included deliberately driving the probe into Jupiter's atmosphere at the conclusion of the mission. Water on Callisto was a much bigger surprise. Could Callisto now be added to the small, but growing, list of potentially fertile worlds within our solar system?

The possibility of a subsurface ocean arises from data on the local magnetic field around Callisto. Callisto does not possess an interior magnetic field but orbits well within the boundary of Jupiter's magnetic field. During multiple flybys, Galileo measured this magnetic field and detected fluxuations in its intensity. The local magnetic environment around Callisto is similar to an electromagnet. Whereas electromagnets have magnetic fields that are induced by the flow of electrons through a looped wire, Jupiter's magnetosphere does the opposite, capturing charged particles from the solar wind and creating electric currents in space. Galileo's instruments showed that this magnetic field was altered by increased conductivity from the satellite itself. While surface ice would not have any effect, the phenomenon could be explained by a subsurface ocean with a salinity level similar to Earth's oceans, conduction of current due to the presence of dissolved salts. This hypothesis is supported by the fact that similar data were taken at Europa, where planetary scientists are more confident that water exists below the surface.

More controversial is the continuing debate over Callisto's differentiation, or lack thereof. This controversy arose from data regarding Callisto's moment of inertia, a measurement of mass that indirectly comes from a body's rotation. This is the phenomenon that controls an ice skater's rotation, increasing it if the arms are brought close to the body and decreasing it when the arms are extended outward. Plane-



The Galileo spacecraft returned data from Callisto that revealed that the Jovian moon may have a salty ocean underlying its icy crust, as shown in this artist's rendering. (NASA/JPL)

tary scientists take this information one step further to determine the composition of a planet or satellite. A moment of inertia of 0.40 would mean that Callisto is totally undifferentiated. Data from multiple passes by Galileo showed a moment of just 0.38, within one standard deviation of theoretical uniformity. This debate is likely to continue for many years, until another spacecraft is sent. Regardless of the answer, the idea that Callisto is not as differentiated as Ganymede, a satellite similar in size and in distance from Jupiter, hints at an interesting beginning of the Jovian system. Answering the question of Callisto's interior will give scientists insight into planet and satellite formation.

While its innards will remain a mystery, Callisto's surface has helped astromoners understand more about comets, comet impacts, and Jupiter's role as protector of the solar system's inner planets (those between it and the Sun). Before the discovery of Comet Shoemaker-Levy 9, the idea of comets impacting planets was not universally accepted. Watching the comet slam into giant Jupiter, and the subsequent "bruises" it temporarily left behind, made the idea of cometary impacts more acceptable. Scientists also learned that it was Jupiter that caused the comet to split into fragments in the first place, leading many to believe that the gas giant has done this in the past. The fact that crater chains exist on the Jupiter-facing hemisphere of Callisto is evidence of past impacts and further evidence that Jupiter is the vacuum cleaner of the solar system, keeping the inner planets safe from dangerous debris.

Finally, studying Callisto may reveal much about the future of humankind, specifically the possibilities of colonizing the solar system. Project HOPE, or Human Outer Planet Exploration, is a futuristic concept mission put forth by NASA. Part of this exploration would include a crewed mission to Jupiter, with a landing on Callisto. Callisto is an optimal choice for a human landing for two reasons. The first is its icy surface, which would provide both a source of water, allowing astronauts to "live off the land," and an opportunity for a first-rate study of impact geology. Second, Callisto's orbit places it in a region of low radiation from Jupiter. This remote, icy outpost would make an excellent location from which to study the Jovian system's past, present, and future.

CONTEXT

Callisto is a wonderful example of how taking a second look leads to a different perception. The Voyager images offered snapshots of Callisto while racing through the solar system's highway. The Galileo probe effectively pulled over and took a look around. Missions like Galileo, which observed the Jovian system from late 1995 to 2003, and Cassini, which began observing Saturn in 2004, offer a chance to understand the distant gas giant planets along with their rocky satellites. Data from Galileo have pointed to the possibility of water on Callisto and have produced debates over its internal structure and its trace of atmosphere—all from a world previously thought dead. Callisto has shown that every object in the solar system has a distinct and complicated personality, arising from a mysterious past, and that we have a long way to go when it comes to understanding our fellow travelers around the Sun.

Michael P. Fitzgerald

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. New York: Cambridge University Press, 2007. A collection of articles provided by recognized experts in their fields of study, this volume offers a comprehensive look at the biggest planet in the solar system. Excellent repository of photography, diagrams, and figures about the Jovian system and the various spacecraft missions that unveiled its secrets.

Carlson, Robert W. "A Tenuous Carbon Dioxide Atmosphere on Jupiter's Moon Callisto." *Science* (February 5, 1999): 283ff. A discussion of Galileo data regarding CO_2 in Callisto's atmosphere.

Cole, Michael D. *Galileo Spacecraft: Mission to Jupiter*. New York: Enslow, 1999. Provides a full description of the Galileo spacecraft, its mission objectives, and science returns through the primary mission. Particularly good at describing mission objectives and goals. Suitable for a younger audience.

Harland, David H. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. New York: Springer Praxis, 2000. Provides virtually all of NASA's press releases and science updates during the first five years of the Galileo mission in a single volume, along with an enormous number of diagrams, tables, lists, and photographs. Also provides a preview of the Cassini mission. Although the book's coverage ends before completion of the Galileo mission, what is missing can easily be found on numerous NASA Web sites.

Khurana, K. K., et al. "Induced Magnetic Fields as Evidence for Subsurface Oceans in Europe and Callisto." *Nature* 395 (October 22, 1998). This article is the resource for all discussions of the possible subsurface ocean on Callisto.

Leutwyler, Kristin, and John R. Casani. *The Moons of Jupiter*. New York: W. W. Norton, 2003. Casani was the original Galileo program manager, and this book offers a heavily illustrated discussion of the Galilean satellites as well as a number of the lesser known Jovian satellites. The authors attempt to accompany their scientific findings with an artful text, which may please the tastes of some readers more than others.

McKinnon, William B. "Mystery of Callisto: Is It Undifferentiated?" *ICARUS* 130 (1997): 540-543. This article explains why the question of Callisto's differentiation does not have a definitive answer.

Melosh, H. J., and P. Schenk. "Split Comets and the Origin of Crater Chains on Ganymede and Callisto." *Nature* 365 (October 21, 1993). Discusses the hypothesis that crater chains on Callisto come from previous comets similar to Shoemaker-Levy 9.

Showman, Adam P., and Renu Malhotra. "Galilean Satellites." *Science* 286 (October 1, 1999). An excellent overview of Jupiter's four largest satellites.

See also: Enceladus; Eris and Dysnomia; Europa; Ganymede; Iapetus; Io; Jovian Planets; Jupiter's Satellites; Lunar Craters; Lunar History; Lunar Interior; Lunar Maria; Lunar Regolith Samples; Lunar Rocks; Lunar Surface Experiments; Mars's Satellites; Miranda; Neptune's Satellites; Planetary Formation; Planetary Satellites; Pluto and Charon; Saturn's Satellites; Titan; Triton; Uranus's Satellites.

Ceres

Category: Small Bodies

Discovered in 1801, Ceres, named for the Roman goddess of agriculture, is the largest of the main-belt asteroids. For a short time it was believed to be the eighth planet in the solar system, but discovery of additional large main-belt asteroids influenced astronomers to revoke its planetary status. Discoveries of Pluto-sized objects in the Kuiper Belt have once again brought Ceres back into the discussion of what constitutes the definition of a planet.

OVERVIEW

On the night of January 1, 1801, the Italian astronomer Giuseppe Piazzi was observing the heavens when he noted a faint object that did not appear on his star charts. At first he thought it might be a comet, but it did not have the typical "fuzzy" appearance associated with comets. If not a comet, what could it be? By observing its motion over the next several weeks, Piazzi was able to determine that its orbital speed was greater than that of Mars but slower than that of Jupiter. This suggested to him that the object must lie between the orbits of Mars and Jupiter.

Additional help came from the German mathematician Carl Friedrich Gauss, who had perfected a means of calculating orbital motion based on limited observations. When he applied his method to Piazzi's observations, Gauss was able to calculate where and when this mysterious object should next appear, and it did just as he predicted. Within one year of Piazzi's discovery, Heinrich Olbers and Franz von Zach were able to relocate Ceres and refine its 4.6 Earth-year orbit. Later scientists were able to determine that it has a spherical shape with a 930-kilometer diameter. In comparison to the Moon, Ceres is one-third its size, but with only less than 2 percent of its mass, giving it a much lower density of 2.1 grams per cubic centimeter.

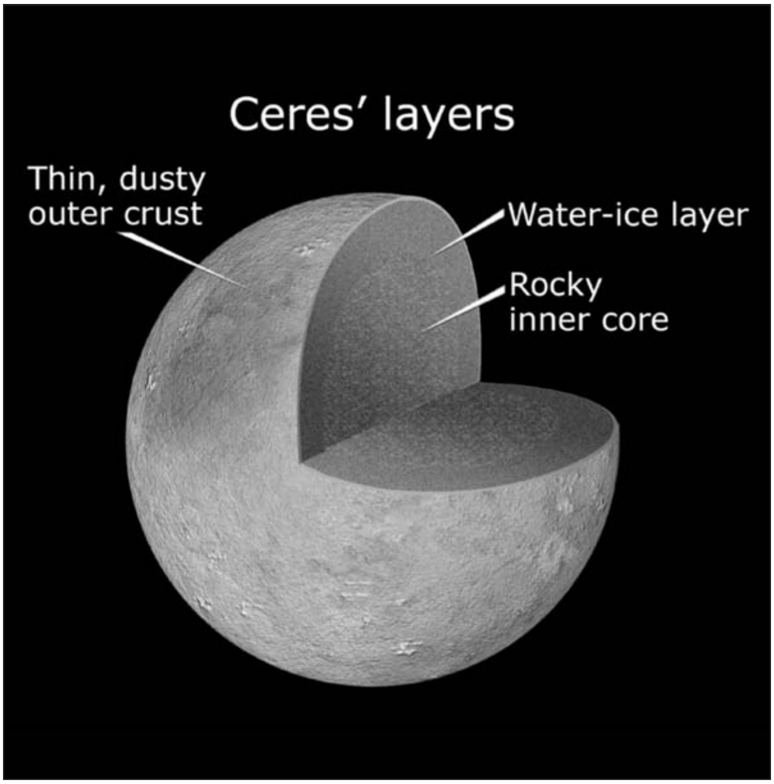
Piazzi's observations and the calculations of Gauss led many of the leading scientists of that time to believe that a new planet had been discovered. This conclusion seemed logical, based on an earlier idea first suggested by Johann Daniel Titius of Wittenberg and later championed by Johann Elert Bode. In 1792, Bode pointed out an apparent mathematical relationship between the distances of the various planets to the Sun. He suggested that the planets were positioned at specific distances from each other based on a mathematical ratio that would later be referred to as Bode's law. This worked reasonably well for all the planets from Mercury through Uranus, with the exception of an apparent gap between Mars and Jupiter. When Piazzi found his mystery object positioned in this gap where Bode suggested a planet should be, this seemed to be the observational confirmation of Bode's law. Even though modern science treats Bode's law as more of an interesting coincidence rather than a scientific law, it did serve a purpose at that time and contributed to the eventual discovery of Neptune.

Although initially proclaimed the eighth planet in 1801, Ceres did not long retain its planetary status. The excitement created in the astronomical community by the discovery of Ceres led to a systematic search of the heavens, which centered on the plane of the ecliptic. Scientists believed that many new and interesting objects would soon be found, and they were right. Within the next six years, three new asteroids—Pellas, Juno, and Vesta—were found within the same general vicinity as Ceres. With four minor bodies now occupying the same region of space, scientists concluded that no one planet would fill the gap in Bode's law. A new theory would have to be created to explain the presence of so many small bodies occupying planetary positions.

Since that time, many theories have been created to explain the presence of the main-belt asteroids. One of the more popular, but incorrect, theories described a large planet exploding and creating a huge number of smaller bodies ranging from the largest, Ceres, down to extremely small meteoroid-sized fragments. Perhaps the most widely accepted theory now describes a "planet that never formed." This theory can be supported by the generally accepted nebular hypothesis of planetary formation, which envisions a final stage of accretion during which a huge number of smaller bodies are attracted to each other and form into a much larger object. In the case of Ceres and the other main-belt asteroids, that final stage was interrupted and they never fully accreted into a single large object.

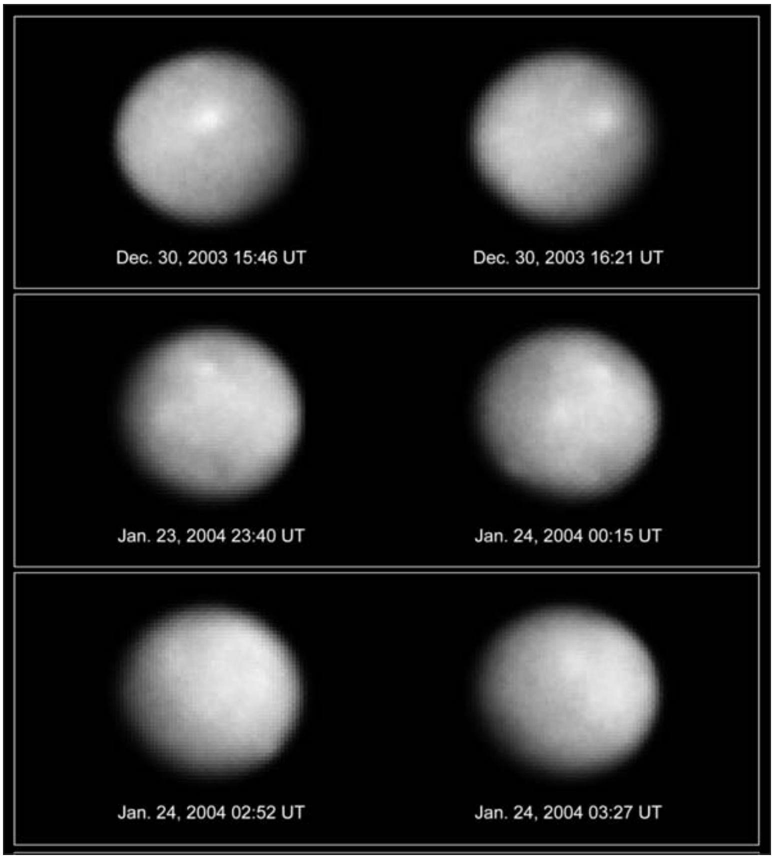
Beginning in the late twentieth century, the study of Ceres and its family of asteroids was no longer regulated by the limitations of Earth-based telescopic observations. The Hubble Space Telescope operating well above the Earth's atmosphere revealed details never before seen by surface-based telescopes. In addition, modern astronomers have the opportunity to send their spacecraft-borne instruments directly to the asteroids to get close-up views of their surfaces and analyze their mineralogical compositions. Several flyby spacecraft missions have investigated a number of smaller asteroids revealing previously unimagined surface conditions. One in particular, the NEAR Shoemaker probe, first orbited and then actually landed on Eros, giving scientists their first detailed images from the surface of an asteroid. The Japanese probe Hayabusa is believed to have landed on the surface of the asteroid Itokawa and collected a sample for return to Earth. Data returned by these missions have rewritten the textbooks on what is known about asteroids. The Dawn spacecraft, en route to both Ceres and Vesta, was designed first to orbit Vesta in 2011 and then leave orbit and go on to rendezvous with and orbit Ceres in 2015. Dawn's mission was to collect sufficient data to help scientists gain a better understanding of the conditions present at the initial accretion stage of planetary formation in the solar system.

Ceres is very different from Vesta. Studies based on a comparison of densities and surface reflectivity have shown that Ceres may have what is considered to be a "wet" surface, composed of water-bearing minerals, as opposed to the "dry" surface minerals of Vesta. Some scientists speculate that Ceres may have a total amount of water locked up in its interior that could rival that of the Earth's surface, while



An artist's conception of the interior of Ceres, formerly classified as an asteroid and now the only dwarf planet in the inner solar system. (NASA/ESA/A. Field, STScI)

Vesta is more comparable to Earth's moon. Other Dawn experiments should shed light on the internal structure, shape, composition, and mass of these two primordial bodies. Scientists have been able to determine that a particular class of meteorites, the HED achondrites, is probably derived from Vesta. It is believed that most meteorites are fragments of crustal material that was blasted off an asteroid's surface during the accretion process or from later impacts. This is based on reflectivity studies of minerals present on the surface of Vesta and from radioisotope chronology studies of the HED meteorites. All evidence seems to point to Vesta. No such data exists for Ceres. This is why a spacecraft mission to Ceres is so important.



Released in 2005, these six images of Ceres were captured by the Hubble Space Telescope between December, 2003, and January, 2004. (NASA/J. Parker et al., Southwest Research Institute)

KNOWLEDGE GAINED

The discovery and later scientific study of the asteroid Ceres had a profound effect both on our early view of the solar system and on our subsequent understanding of the origin and nature of planetary bodies. At the time of its discovery (1801), Ceres represented another "new" object in the heavens that was unknown to the ancient astronomers. It had been less than fifty years since the return of Halley's comet (1758) had literally galvanized the concept of gravity and its effects on motion, and only twenty years after the discovery of Uranus (1791). The discovery of three other asteroids would soon follow, but after 1807, no other asteroid was detected until 1845.

The following year Neptune was discovered, and astronomers had definitely moved into the "modern era" with bigger and better telescopes, technology that included spectroscopy and photography, and a more scientific perspective of the universe. With Neptune recognized as the eighth planet, the asteroids fell into their proper place within the structure of the solar system. The study of asteroids remained within the domain of observational astronomers until scientists and engineers could develop the technology to send their scientific instruments to the planets and minor bodies. Once this happened, the outpouring of data changed our view of the nature and origin of the planets.

To the early astronomers Ceres was only a faint speck of light in the night sky. Today modern astronomers see it quite differently, primarily as a result of the observations made by the Hubble Space Telescope's Advanced Camera for Surveys. In 2005 astronomers observed Ceres through a complete nine-hour revolution, taking 267 photographic images. From these observations astronomers were able to determine that Ceres has a spherical shape with a diameter slightly wider at the equator than at the poles. This suggests that it has a differentiated internal structure with denser materials forming a core and lighter materials closer to the surface. Scientists suspect that, because Ceres' density is much lower than Earth's, large amounts of water ice may exist either on the surface of Ceres or buried within its crust. This surmise is supported by spectral evidence for water-bearing minerals that may be present on the surface that are not representative of Ceres' crystal rocks. Additional microwave studies suggest that this surface material might be dry clay. All considered, Ceres could turn out to be the most Earth-like body in the solar system; it may even be a haven for primitive forms of life.

CONTEXT

Clearly the asteroids hold many vital clues to unraveling the mysteries surrounding the formation of the planets. On the basis of their respective sizes, densities, and chemical compositions, the planets in our solar system are divided into three major groups: the terrestrial (Earth-like) planets, the Jovian (Jupiter-like) planets, and the dwarf planets. The third group can be further divided into rocky objects like Ceres and into ice bodies like Pluto. These dwarf planets, rocky or ice, most likely represent a fundamental primordial stage in the formation of planets.

By studying these early remnants of planetary formation, scientists can achieve a clearer picture of their formative processes. Each group will have its own distinctive secrets to reveal. The rocky dwarf planets positioned between Mars and Jupiter formed under conditions of higher temperature, higher density, and higher velocity than the icy worlds at the edge of the solar system in the Kuiper Belt. It is believed that at this distance from the Sun, these objects have remained essentially unchanged over the last 4.6 billion years. In 2015, the New Horizons spacecraft will visit Pluto and send back images and data giving science its first close-up look at this unknown world. In that same year the Dawn spacecraft will orbit Ceres. Perhaps then, with both worlds under study, science will be able to fill in many of the gaps in our understanding of planetary formation.

Paul P. Sipiera

FURTHER READING

Bell, Jim, and Jacqueline Mitton, eds. *Asteroid Rendezvous: NEAR Shoemaker's Adventures at Eros*. Cambridge, England: Cambridge University Press, 2002. A collection of nine scientific articles that provide the reader with an overview of an asteroid rendezvous mission and what to expect from the anticipated Dawn mission to Ceres and Vesta. Suitable for a wide range of readers.

Bottke, William F., Jr., Albertoi Cellino, Paolo Paolicchi, and Richard P. Binzel, eds. *Asteroids III*. Tucson: University of Arizona Press, 2002. This comprehensive work is a compilation of scientific papers that cover virtually every aspect of asteroid research. The paper on Piazzi and Ceres is especially relevant. Best suited for the graduate student and professional scientist.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An excellent reference for a variety of topics in planetary science, including asteroids. Suitable for readers of high school-level and above.

Kowal, Charles T. *Asteroids: Their Nature and Utilization*. Chichester, England: Ellis Harwood, 1988. A good basic reference source for general information about asteroids, including both their potential for commercial use and their potential as a hazard to Earth. Suitable for both astronomy enthusiasts and students at the undergraduate and graduate levels.

Lang, Kenneth R. *The Cambridge Guide to the Solar System*. Cambridge, England: Cambridge University Press, 2003. A concise yet comprehensive book, containing a wealth of information on the members of the solar system. Excellent for a wide range of readers.

Reedy, Francis. "The Tenth Planet." *Astronomy* 33, no. 11 (2005): 68-69. A good basic article for a general readership, describing the scientific controversy over the definition of what constitutes a planet.

Sipiera, P. P. "Dawn Mission." In *USA in Space*. 3d ed. Edited by Russell R. Tobias and David G. Fisher. Pasadena, Calif.: Salem Press, 2006. A concise yet comprehensive article describing the current space mission to the as-

teroids Ceres and Vesta. Suitable for a wide range of readers.

See also: Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Kuiper Belt; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Planetary Classifications; Pluto and Charon.

Comet Halley

Category: Small Bodies

Halley's comet is the brightest, most famous of the known periodic comets. Definitive records of sightings go back more than two thousand years. The comet travels around the Sun roughly once every seventy-six years in a highly eccentric retrograde orbit inclined 20° to the ecliptic plane. Its orbital period has enabled many observers to see Halley's comet twice during their lifetimes.

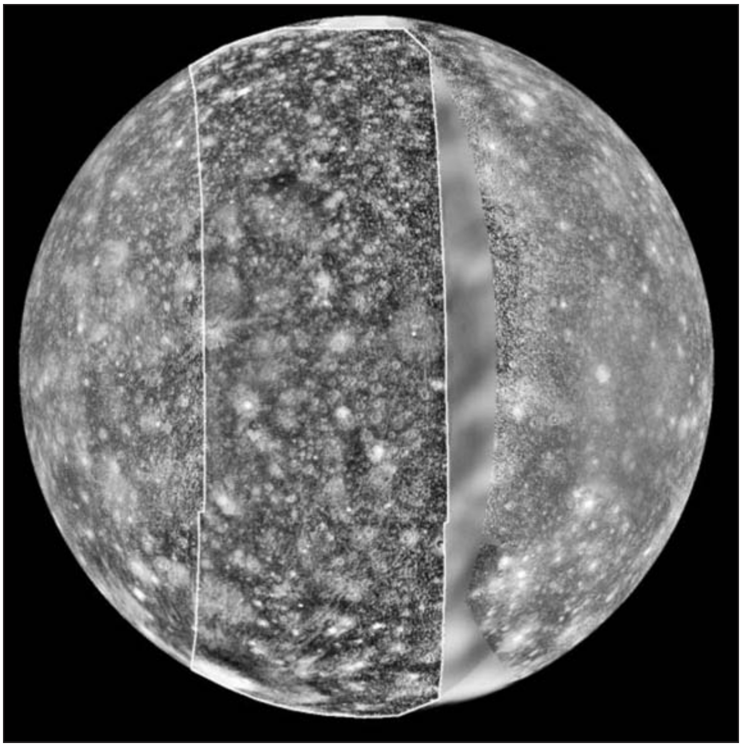
OVERVIEW

For many years, the idea that comets were "dirty snowballs" has generally been accepted by astronomers. First proposed by Fred L. Whipple in 1950, this was one of a number of different ideas about the makeup of comets. The most popular idea was that they were "flying sandbanks," or collections of interstellar dust and gas accreted as the Sun and planets periodically passed through vast clouds of interstellar matter in their journey through the galaxy. The Sun's gravity then drew in the material that eventually collected to form individual bodies. This idea was popular during the first half of the twentieth century and was championed by British astronomers R. A. Lyttleton and Fred Hoyle. Since the middle of the nineteenth century, meteor streams have been associated with comets, and supporters of the "flying sandbank" model of cometary nuclei suggested that the particles within meteor streams arose from material escaping from comets as they moved through the solar system.

It is now widely believed that cometary nuclei are composed of material that condensed from the solar nebula at the same time as did the Sun and its planets. The European Space Agency's (ESA's) and other spacecraft that intercepted and studied Halley's comet in March, 1986, detected copious amounts of carbon, nitrogen, and oxygen. The materials given off by the comet signify that these objects were formed in the outer regions of the solar system, where the extremely low temperatures necessary for them to solidify prevailed. Giotto revealed that the nucleus of Halley's comet is a tiny, irregularly shaped chunk of ice coated by a layer of very dark material measuring some 15 kilometers long by 8 kilometers wide. This layer is thought to be composed of carbon-rich compounds and has a very low albedo, reflecting merely 4 percent of incident light. This low reflectivity makes the nucleus of Halley's comet one of the darkest objects known. However, various bright spots were seen on the nucleus. A hill-type feature was found near the terminator, along with one resembling a crater located near a line of vents. The vents seen on the nucleus appear to be fairly long-lived. Dust jets detected by the Russian Vega 1 and Vega 2 probes appear to have emanated from these vents, two of which were also identified by Giotto. Possibly some of the larger vents have survived successive perihelion passages.

Gas and dust that cause all the cometary activity seen (including the coma and tail) emanate from the nucleus via localized vents or fissures in the outer dust layer. These vents cover approximately 10 percent of the total surface area of the nucleus. They become active when exposed to the Sun and cease to expel material when plunged into darkness as the nucleus rotates. The force of these jets of material escaping from the nucleus plays an important role in the comet's motion around the Sun, affecting its orbital speed. Halley's comet was several days late in reaching perihelion during the last apparition in 1986, a result of the jetlike effects of the matter being expelled, as a consequence of Newton's second law of motion. The late arrival of

Halley's comet was one of the factors examined by Swedish astronomer Hans Rickman, who attempted to calculate the mass of the nucleus from the amount of ejected material. Linking the ejection rate to the delay in perihelion, he judged the volume of the nucleus to be between 50 and 130 cubic kilometers. Measurements obtained through spacecraft imagery, however, revealed a volume closer to 500 cubic kilometers. The only conclusion was that the nucleus is markedly porous and far less dense than first anticipated, with an average density of no more than a quarter that of ice. This porosity meshes with the belief that comets formed in the outer regions of the solar nebula, where material coming together would remain loosely bound rather than compacting.



Comet Halley's nucleus. (NASA/GSFC)

The fact that the nucleus of Halley's comet rotates is not in doubt. What does remain unresolved is the period of rotation. Using photographs of the comet taken during its apparition in 1910, astronomers calculated the rotation period to be 2.2 days around an axis that was fairly well aligned with the poles of the comet's orbit around the Sun. Results obtained by the Giotto, Vega, and Japanese Suisei probes appeared to support this value. Ground-based observations carried out during 1986, however, indicated a rotation period of 7.4 days. This value was supported by other ground-based observations, together with results from the American Pioneer Venus orbiter, which examined Halley's comet when it neared perihelion. Controversy ensued over these differing values, although a possible explanation has been suggested. The nucleus of Halley's comet could actually display both periods of rotation; one being spinning around its axis and another the precession of the axis of rotation. The combination of rotation and precession is still contested by certain astronomers, to some extent because of the porosity of the nucleus. Any precessional properties would quickly disappear unless the nucleus were fairly rigid.

Comets give off copious amounts of gas and dust that spread out in tails across large areas of space. Investigation of this material can reveal much about the composition of cometary interiors. Many of the investigations carried out by the European, Japanese, and Soviet space probes were directed toward a survey of the material ejected by Halley's comet. These investigations were supplemented by observations both from ground-based astronomers and from the American Pioneer Venus, International Cometary Explorer (ICE), and International Ultraviolet Explorer (IUE) spacecraft. As is the case with the surface of Halley's nucleus, the dust thrown off by the comet was found to be very dark and may have emanated from the surface itself rather than the interior. Giotto and Vega carried out analyses of the dust. They found a mixture of different materials, including the lighter elements oxygen, hydrogen, nitrogen, and carbon, and the heavier elements silicon, iron, and magnesium. The amount of carbon found during the investigations coincides quite well with the observed abundance of this material elsewhere in the galaxy, indicating that comets are made of interstellar material.

More than three-quarters of the gas ejected from the nucleus was found to be water vapor, which also appears to constitute more than 80 percent of the nucleus. The rate of production varied during the interval the comet was examined by these space probes. Vega 2 found approximately 16 tons of water coming away from the nucleus during its flyby, while Vega 1 detected double that rate. These large changes are reflected in the fact that Comet Halley's brightness sometimes varied by a factor of two or three from night to night. The velocity of the ejected vapor was found to be between 0.8 and 1.4 kilometers per second. This was the first time that water had been positively identified in a comet, in spite of the fact that cometary nuclei were widely thought to consist of a mixture of dust and water ice. Carbon monoxide and carbon dioxide were also detected, although methane was not found at all. This is strange, in that either any methane which existed in the comet may have been altered chemically during the period since the formation of the comet or methane was lacking in the cloud of material from which the comet formed. If there is methane in Halley's nucleus, it must constitute a very tiny percentage of the total makeup.

Processes involved in the release of gas from the nucleus may have played a prominent role in the evolution of its surface. It has been suggested that, as a comet approaches the Sun after spending its time in temperatures of approximately 40 kelvins in the outer regions of the Sun's influence, warming effects of the Sun can cause the ice within the nucleus to expand. This would result in heat generation and release of trapped gas. Some of this gas may collect in pockets, which eventually explode, producing craterlike features similar to that imaged by Giotto.

METHODS OF STUDY

Halley's comet is unusual (though not unique) in that it is named for the astronomer who first calculated its orbital path rather than the person who discovered it. Edmond Halley observed a bright comet in 1682, the impression of this sighting staying with him and eventually expanding into a deeper interest in comets. In 1705, Halley began a study of a number of bright comets seen between 1337 and 1698. Using methods developed by Sir Isaac Newton, he carried out work on the orbital motions of some twenty-four comets seen during this period. He noticed from his results that there were many similarities between the orbits of the comets observed in 1531 and 1607 and the bright comet he had seen in 1682. The intervals between the sightings were also roughly identical at around seventy-six years. This led Halley to predict that these sightings were of the same comet, and that it would reappear in 1758.

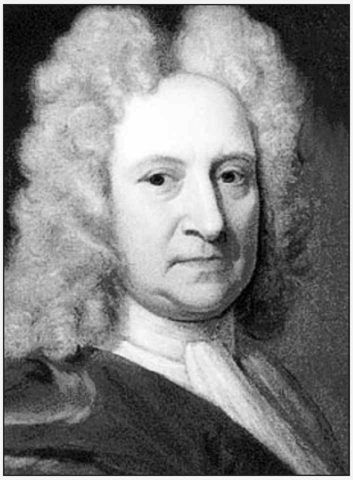
Halley died in 1743, although astronomers began a search for the returning comet as the date forecast by Halley drew near. French astronomer and mathematician Alexis-Claude Clairaut, with the help of Joseph-Jérome de Lalande and Madame Nicole Lepaute, attempted to calculate its orbital path in more detail. Taking into account gravitational effects of Jupiter and Saturn, they calculated that the comet would reach perihelion on April 13, 1759, and published ephemerides (detailed star maps and charts) to help astronomers with their search. Many famous astronomers joined in, although it was the amateur astronomer Johann Georg Palitzsh from Dresden who first spotted the comet on Christmas Day, 1758. The reappearance was quickly confirmed, and the comet was named for Halley in honor of the fact that he had correctly predicted its return. Once a number of observations had been obtained, a revised orbit was calculated. It was found that Clairaut's calculated perihelion date was in error by thirty-two days. Scientists were at a loss to explain this error, although they did not know about the existence of the two giant planets Uranus and Neptune, which were not to be discovered until 1781 and 1846, respectively.

Since the 1758 appearance, Halley's comet has been seen on three occasions: in 1835, in 1910, and in 1985-1986. Times of previous visits of the comet have been calculated by taking into account the gravitational effects of other bodies of the solar system and plotting the comet's

The Adventurous Edmond Halley

In 1684, Edmond Halley was a young scientist who had already made a name for himself as a precocious astronomer: He was the first to observe that the Sun rotated on an axis, during a trip to St. Helena in the South Seas. In 1680, during his Grand Tour of Italy and France, he had observed the comet that would bear his name. He had produced star catalogs and tidal tables, and he was trying to determine why Kepler's laws worked the way they did. Then, in April, his father's disfigured corpse was discovered near a riverbank; he had been missing for more than a month. Edmond's attention was redirected toward a bitter battle with his stepmother over the family estate.

Four months later, Halley was visiting Isaac Newton, who had solved the problems with Kepler's laws but had "misplaced" the solutions, supposedly worked out when Cambridge had been shut down during the plague of 1665. Halley began a campaign of diplomacy to get the eccentric and overly sensitive Newton to publish his results before someone else (Robert Hooke) derived the inverse square law and beat him to it. This was the genesis of Newton's *Principia* of 1687, published at Halley's expense.


(NASA)

In the meantime, Halley was supporting himself as a clerk at the Royal Society and working on a diverse array of projects, from determining the causes of the biblical Flood (which he unorthodoxly and dangerously placed earlier than the accepted date of 4004 B.C.E.) to making the connection between barometric pressure and altitude above sea level. He even calculated the height of the atmosphere, at a remarkably accurate 45 miles. Motivated by his persistent lack of money, Halley also designed various nautical instruments: a prototype diving bell, a device for measuring the path of a ship, and another device for measuring the rate of evaporation of seawater. He even prepared life-expectancy tables that became the basis for modern life insurance. Between 1696 and 1698, he became the deputy comptroller of the Royal Mint at Chester, a post offered him by Newton, who was then the warden of the Mint. Administration did not prove to be one of Halley's many talents, however, and Newton found himself having to defend his friend against the Lord's Commissioners.

In 1698, Halley set out on another expedition to the South Seas to study the magnetic variations of the Earth's compass. The journey was abandoned (with the ship's first lieutenant facing a court-martial on their return), but Halley tried again a year later with more success. He also went on a secret mission in 1701, about which little is known, traveling to France for the Admiralty on the pretext of yet another scientific expedition. In 1703, Halley became a member of the Council of the Royal Society in recognition of his work, and in the same year, he was appointed to the Savilian Chair of Geometry at Oxford, where he conducted his study of comets. It was around this time that he made the observation for which he became famous:

> Many considerations incline me to believe that the comet of 1531 observed by Apianus is the same as that observed by Kepler and Longomontanus in 1607 and which I observed in 1682. . . . I would venture confidently to predict its return, namely in the year 1758. If this occurs there will be no further reason to doubt that other comets ought to return also.

In 1719, on the death of John Flamsteed, Halley succeeded to the post of Astronomer Royal, a position he held until his death in 1742. The practicality and range of his interests made him a celebrity whose achievements far exceeded those for which he is remembered today. He did not live to see his comet, which was sighted on Christmas, 1758.

orbital course backward through time. Dates calculated for previous apparitions have been substantiated by checking against ancient astronomical records, primarily those of Chinese astronomers. The first definite appearance of Halley's comet took place in 240 B.C.E., although the 12 B.C.E. appearance is the first about which detailed information is available. The most famous return was that of 1066, which was interpreted as a bad omen by the Saxons and, in particular, by Harold, the last of the Saxon kings. William of Normandy, who viewed the apparition as a good sign, invaded England, following which Harold died at the Battle of Hastings in October of that year. The Bayeux Tapestry depicts the comet suspended above Harold, who is seen tottering on his throne as his courtiers look on in awe and terror.

The 1531 appearance is important because it was one of two apparitions studied by Halley (the other being that of 1607) prior to his two deductions: that these historical sightings were of the same object, and that the comet is a regular visitor to this region of the solar system. A comprehensive set of observations of the 1531 appearance was made by astronomer Peter Apian, who published his results in 1540. The 1607 appearance was observed and recorded by many astronomers, including Johannes Kepler. This was the last apparition of Halley's comet before the introduction of the telescope.

After the comet's reappearance in 1758 and the discovery of Uranus in 1781, astronomers were able to plot its orbit with even greater accuracy. Long before its scheduled return in 1835, many attempts were made to calculate the expected date of perihelion passage. The consensus was that Halley's comet would pass closest to the Sun in November, 1835. The search for the returning comet started as early as December, 1834, almost a year before it was due to sweep through the inner solar system. The first sighting was not made, however, until August 6, 1835, by Father Dumouchel and Francisco di Vico at the Collegio Romano Observatory. Confirmation came via Friedrich Georg Wilhelm von Struve, who saw the comet on August 21. Perihelion occurred on November 16.

Prominent among the astronomers who studied the comet during the 1835 apparition was Sir John Frederick Herschel, who was then based at a temporary observatory near Cape Town, South Africa. He was in the process of completing the sky survey started by his father, Sir William Herschel, and had moved to South Africa in order to survey the southern stars that were visually inaccessible from England. John Herschel made his first attempt to locate the comet in late January, 1835, although he did not see it until October 28. The 1835 apparition was remarkable in that much activity was seen to occur in the comet. Prior to its temporary disappearance in the Sun's rays as it rounded the Sun, a number of changes were observed in the tail. These disturbances continued after its reappearance. The tail was seen to vary noticeably in length. The head also altered in appearance, at times appearing almost as a point of light, while at others taking on a nebulous form. It was noticed that the coma expanded while undergoing a reduction in brightness, eventually becoming so dim that it merged into the surrounding darkness. Herschel's final observation of Halley's comet in mid-May, 1836, was the last that any astronomer made until the 1910 return. All data scientists have about the 1835 apparition are in the form of sketches and visual descriptions. Photography had not yet made an impact on astronomy, although the appearance in 1910, through the use of the camera, provided the most comprehensive and detailed study of Halley's comet up to that time.

The third predicted return in 1910 was awaited eagerly by astronomers all over the world. The interval between the 1835 and 1910 visits had been littered with numerous bright comets, notable among which were the Great Comet of 1843, Donati's comet of 1858, and the Great September Comet of 1882. The latter is particularly significant in that it was the subject of the first successful attempt to photograph a comet. A good image was obtained by Sir David Gill in South Africa. Observation of Comet Morehouse in 1908 demonstrated that a series of photographs was an ideal means of monitoring cometary structural changes. (Comet Morehouse itself underwent a number of prominent changes that, coupled with the fact that Halley's comet had suffered in a similar fashion three-quarters of a century before, whetted the appetites of astronomers who were gearing up for the forthcoming apparition.) The prolonged period of cometary activity following its last visit had allowed astronomers to perfect their observing techniques and paved the way for observations of the return of Halley's comet.

The comet had passed aphelion in 1872, after which it once more began its long journey toward the inner solar system. The first astronomer to detect the returning visitor was astrophysics professor Max Wolf at Heidelberg, Germany. A photographic plate was exposed on the night of September 11-12 and recorded the comet close to its expected position. It did not become visible to the naked eye until well into 1910. Prior to this, another bright comet made an unexpected appearance. The Great Daylight Comet was first spotted by diamond miners in Transvaal, South Africa, in the early morning sky on January 13, 1910. Confirmation of the discovery was made four days later, and news of this spectacular discovery was distributed to the world's observatories. Unlike Halley's comet, which was to appear later that year, the Great Daylight Comet became a brilliant evening object for observers in the Northern Hemisphere. Its tail attained a maximum length of 30º or more by the end of January. The comet became so bright that it was visible to the naked eye even in broad daylight (hence its name).

The Great Daylight Comet was widely mistaken for Halley's comet by many people who had been expecting its return at about this time, although Halley's comet did not put on as grand a show. Bad weather together with the fact that a full moon occurred at what should have been the best time for observation meant that astronomers north of the equator were disappointed. Yet, even working against these odds, they did obtain many useful photographs and were able to study the comet spectroscopically. The best results, however, were obtained from observatories in the Southern Hemisphere, notably at Santiago in Chile. From mid-April to mid-May, 1910, Halley's comet was in the same area of the morning sky as Venus, the two objects together forming a marvelous visual spectacle in the constellation of Pisces. Much activity was noted in both the nucleus and the tail of the comet. Sequences of photographs showed marked changes in the head, including material being ejected from the nucleus and halos expanding out from the nucleus. The tail also underwent violent changes, with material being seen to condense in various regions. On April 21, the day following perihelion, the previously smooth northern edge of the tail became irregular and distorted. Material seemed to be thrown out in various directions, and parts of the tail seemed to be ejected into space, an event clearly visible on photographs obtained at the time. For some days following perihelion, a jet of material from the nucleus seemed to be refueling the northern section of the tail. Once this activity ceased, the tail's southern section increased in brightness. A few weeks after perihelion, the two types of cometary tail appeared, a straight and distinct gas tail contrasting with the fainter, more diffuse and curved dust tail. Halley's comet passed between the Sun and Earth on May 18, although in spite of many attempted observations, no trace of the nucleus could be seen as the comet transited the solar disk. This proved that the nucleus must be tiny and the gas around it very tenuous. During this time, it was thought that the Earth may pass through the tail, although there is no evidence that this actually occurred. The pronounced curve of the tail seems to have taken it away from the Earth, preventing a passage of the planet through it. The closest approach of the comet to Earth was on May 20, when the distance between the two bodies was 21 million kilometers. For a time afterward, the comet became a prominent evening object for American observers, and many useful results were obtained by astronomers at Lick Observatory and Mount Wilson Observatory in California. A number of changes in the comet's structure were seen, and many spectroscopic observations were taken. These showed the presence of a large number of different molecules in the comet, and helped astronomers to understand more clearly its chemical constitution.

As the comet started on its journey back to the outer regions of the solar system, it grew steadily fainter. It was last seen when beyond the orbit of Jupiter in a photograph taken on June 15, 1915, on its way toward aphelion in 1948. The next return would be accompanied by an unprecedented campaign by astronomers and space scientists to expand their understanding of comets in general, and Halley's comet in particular.

The return of 1985-1986, the most recent to date, provided astronomers with their best chance yet of exploring a comet. Unlike other bright comets, many of which appear suddenly, the orbital path of Halley's comet is known with both great precision and great accuracy. Therefore, it was possible to plan missions by robotic space probes to rendezvous with the comet during its last return. For a comet rendezvous mission, the position of the comet at time of interception must be known well in advance, as was the case with Halley's comet. In all, five space probes were sent to examine the comet.

Two of these were the Soviet Vega probes, launched in December, 1984, to release balloons into the Venusian atmosphere. Along their way, the probes encountered Halley's comet on March 6 and March 9, 1986, at distances of 8,890 kilometers and 8,030 kilometers, respectively. Among the equipment they carried were cameras, infrared spectrometers, and dust-impact detectors.

The two Japanese probes carried out their investigations from greater distances. Sakigake, launched in January, 1985, flew by the comet on March 11, 1986, at a distance of 6.9 million kilometers. Its primary purpose was to investigate the interaction between the solar wind and the comet at a large distance from the comet. One of the main aims of Suisei, launched in August, 1985, was to investigate the growth and decay of the hydrogen corona. Suisei flew past the comet on March 8, 1986, at a distance of 151,000 kilometers.

By far the most ambitious, and most successful, of the probes dispatched to Halley's comet was the European Giotto, named in honor of the Italian painter Giotto di Bondone. It launched toward the comet on July 2, 1985. Giotto was cylindrical in shape, with a length of 2.85 meters and a diameter of 1.86 meters. Its payload included numerous dust-impact detectors, a camera for imaging the nucleus and inner coma of Halley's comet, and a photopolarimeter for measuring the brightness of the coma. Giotto flew within 610 kilometers of the nucleus on March 14, 1986, at a speed of more than 65 kilometers per second. Data collected by Giotto were immediately transmitted back to Earth via a special high-gain antenna mounted on the end of the space probe facing away from the comet. Information was received back on Earth by the 64-meter antenna at the Parkes ground station in Australia. At the opposite end, Giotto was equipped with a special shield to protect it from impacts by dust particles during its passage through the comet's halo.

Exploration of Halley's comet by space probes was a truly international effort, the images and measurements obtained by the Soviet Vega craft helping scientists to target Giotto precisely. From Earth, the nucleus of a comet is hidden from view by the material surrounding it. Not until the Vega images were received was its position established and the subsequent trajectory of Giotto determined. During the close encounter, all instruments performed well, although disaster struck immediately before closest approach to the nucleus. A dust particle weighing merely one gram impacted Giotto. This temporarily knocked the spacecraft and its antenna out of alignment with Earth, and for thirty tense minutes contact was lost. The problem was rectified, and contact was reestablished. After the encounter, it was found that

Launched in 1985, the Giotto space probe passed by Comet Halley's nucleus in 1986. (European Space Agency) approximately half of the scientific equipment had suffered damage, although scientists were able to redirect the craft and put it on a course back to Earth. Tests carried out by the European Space Agency in 1989 paved the way for reactivation of Giotto, which set up a pass within 22,000 kilometers of the Earth and placed Giotto in a new orbit that allow it to intercept another comet. On July 10, 1992, Giotto flew close to Comet Grigg-Skjellerup, at a point just twelve days in advance of the comet's closest passage to the Sun, a time when its activity was approaching maximum.

CONTEXT

Although study of Halley's comet has taught scientists much about comets in general, there still remains much to learn about these ghostly visitors. Halley's comet provides a chance to investigate the origins of the solar system. Cometary explorations by space probes could include rendezvous missions during which a probe would position itself close to a cometary nucleus for a prolonged period and perhaps send a lander to the surface of the nucleus. The possibilities of such a mission were being examined by the National Aeronautics and Space Administration (NASA) at the time of Halley's 1986 visit. Known as Comet Rendezvous and Asteroid Flyby (CRAF), this mission would have enabled scientists to undertake close-up exploration of both asteroids and comets. Unfortunately, budget cuts led to the cancellation of CRAF. Sample return missions, by which scientists can examine at first hand material plucked from the heart of a comet, also remain a possibility. Some astronomers and scientists hope for a mission that will carry a human crew to Halley's comet during its next apparition, in 2061.

More realistically in the meantime, NASA was able to launch its Deep Space 1 probe and demonstrate the capability of an ion propulsion system to drive a spacecraft to effect rendezvous with an asteroid and a comet. On September 22, 2001, Deep Space 1 flew within 2,200 kilometers of Comet Borrelly, performing measurements and taking high-resolution images. The Deep Impact mission slammed a copper impactor into the Comet Tempel 1 on July 4, 2005, to expel surface material and excavate a crater on the comet's nucleus. The flyby portion of the Deep Impact spacecraft observed the collision of its impactor and analyzed material thrown up from the formation of an impact crater on the cometary nucleus. Then in January, 2006, the Stardust mission returned samples to Earth released from Comet Wild 2; those samples were collected at a distance of 240 kilometers from the comet's nucleus. ESA launched the Rosetta spacecraft in 2004 and set it on a trajectory toward an encounter with the comet 67P/Churyumov-Gerasimenko in May, 2014. Rosetta is designed to orbit the comet and later release a small lander named Philae to touch down on the comet's nucleus and perform in situ analyses of surface materials.

Brian Jones

FURTHER READING

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. Filled with color diagrams and photographs, this popular work covers solar system astronomy and planetary exploration through the Mars Pathfinder and Galileo missions. Accessible to the astronomy enthusiast. Provokes excitement in the general reader, who gains an explanation of the need for greater understanding of the universe.

Faure, Gunter, and Teresa M. Mensing. *Introduction to Planetary Science: The Geological Perspective*. New York: Springer, 2007. Designed for college students majoring in Earth sciences, this textbook provides an application of general principles and subject material to bodies throughout the solar system. Excellent for learning comparative planetology.

Gingerich, Owen. "Newton, Halley, and the Comet." *Sky and Telescope* 71 (March, 1986): 230-232. Provides background information on Sir Isaac Newton and Edmond Halley. Describes how Halley used Newton's work with gravitation to draw comparisons between the orbital motions of comets seen in 1531, 1607, and 1682, which led to his conclusion that they were all sightings of the same object and his predict of its return in 1758. Suitable for the general reader.

Grewing, M., F. Praderie, and R. Reinhard, eds. *Exploration of Halley's Comet*. New York: Springer, 1989. A technical review of the information garnered by the return of Halley's comet in 1985-1986.

Harpur, Brian, and Laurence Anslow. *The Official Halley's Comet Project Book*. London: Hodder and Stoughton, 1985. A comprehensive guide to knowledge of Halley's comet prior to its exploration by space probe. As well as a general description of comets, the book contains details of Edmond Halley and his work, many facts relating to Halley's comet and its previous appearances, and a detailed description of the 1910 apparition of the comet. Includes a discussion on the pronunciation of Halley's name and a collection of poems written about the comet in 1910. A useful book for the general reader, containing many items not printed elsewhere.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar system astronomy suitable for an introductory college course, filled with supplemental learning aids and solved student exercises. Accessible to nonscientists as well. A Web site is available for educator support.

Sekanina, Zdenek, ed. *The Comet Halley Archive Summary Volume*. Pasadena, Calif.: Jet Propulsion Laboratory (International Halley Watch), California Institute of Technology, 1991. A collection of observations made by various observers of Halley's comet's return in 1985-1986.

Whipple, Fred L. "The Black Heart of Comet Halley." *Sky and Telescope* 73 (March, 1987): 242-245. An examination of the information received regarding the nucleus of Halley's comet and what it tells scientists. Comparisons are drawn between previous models of the structure of cometary nuclei and current knowledge. Suitable for the general reader.

_______. *The Mystery of Comets*. Washington, D.C.: Smithsonian Institution Press, 1985. Chapter 4, "Halley and His Comet," outlines the life and work of Edmond Halley and his involvement with cometary orbits. Chapter 5, "The Returns of Halley's Comet," describes the apparitions of Halley's comet from the earliest sightings to 1910; chapter 24, "Space Missions to Comets," is a description of the various space probes that intercepted Halley's comet during its return in 1985-1986. Suitable for the general reader.

See also: Ceres; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Earth's Oceans; Eris and Dysnomia; Extraterrestrial Life in the Solar System; Infrared Astronomy; Kuiper Belt; Lunar Craters; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Planetary Classifications; Planetary Formation; Pluto and Charon; Solar System: Element Distribution; Solar System: Origins; Solar Wind; Ultraviolet Astronomy.

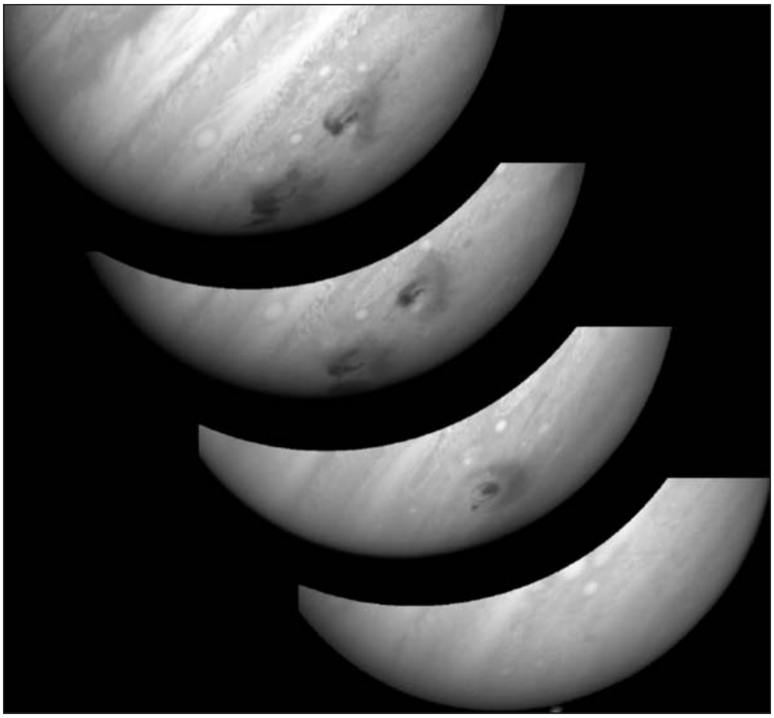
Comet Shoemaker-Levy 9

Category: Small Bodies

The spectacular collision of comet Shoemaker-Levy 9 with Jupiter in July, 1994, provided valuable information about comets, Jupiter's atmosphere, and the Jovian role in diminishing potentially cataclysmic Earth-damaging debris in the inner solar system.

OVERVIEW

On the night of March 24, 1993, the husband-and-wife team of Eugene ("Gene") and Carolyn Shoemaker together with their colleague David Levy were using the Schmidt telescope at Mount Palomar Observatory in California to take photographs in connection with a project designed to discover near-Earth celestial objects. Gene Shoemaker, who had recently retired from the U.S. Geological Survey's Astrogeology Research Program, which he had established, was an expert in Earth-orbit-crossing asteroids and comets. In their five years together, the Shoemaker-Levy team had already discovered eight comets, and they were pleased



A photo mosaic showing the comet Shoemaker-Levy 9 as it impacted Jupiter in 1994. (JPL/NASA/STScI)

when one of their photographs of the sky near Jupiter revealed what Levy called "the strangest comet" he had ever seen. Its several tails and bat-shaped wings of dust reminded him of the American Stealth Bomber. They quickly realized that this fragmented comet was an important discovery, and, three days later, they made comet Shoemaker-Levy 9 (SL 9) public in a circular published by the International Astronomical Union (IAU). Following tradition, the comet was named after its discoverers, with the number indicating it was the ninth comet that this team had found. Its formal IAU name became D/1993F2, in which the prefix "D" indicated that it was a periodic comet that later "disappeared," 1993 was the year of discovery, the suffix "F" represented the half-month of discovery (F = March 16-31), and "2" meant it was the second discovery in that half-month.

Other observers, stimulated by this announcement, returned to photographs that they had taken before March 24 and confirmed the discovery of the Palomar team. Using these and other data, astronomers calculated the orbit of the comet, which, unlike all earlier comets, orbited Jupiter rather than the Sun. Its highly elliptical orbit had an apojove (the point farthest away from Jupiter) of nearly 50 million kilometers and a period of nearly two years. Later data helped refine this orbit and provided clues about the comet's early history. Like most comets, it had orbited the Sun, but several decades earlier it had been captured by Jupiter's gravity. In early July, 1992, it traveled so close to Jupiter, about 20,000 kilometers from the Jovian cloud tops, that the giant planet's powerful gravity broke the comet into twenty-one separate fragments, each of which collected a coma of dust.

Following the convention established for previous fragmented comets, the twenty-one discernible pieces were labeled with letters of the alphabet (excluding I and O), and so SL 9's fragments, which averaged a few kilometers in diameter, ran from A to W, with the brightest (and presumably the largest) piece called Q. With the orbit and fragments identified, astronomers soon realized that this piecemeal comet was on a collision course with Jupiter. For SL 9 there would be no escape from Jovian orbit to return to the Kuiper Belt, now believed to be the source of Jupiter Family (JF) comets; instead, SL 9 faced extinction within sixteen months of its discovery.

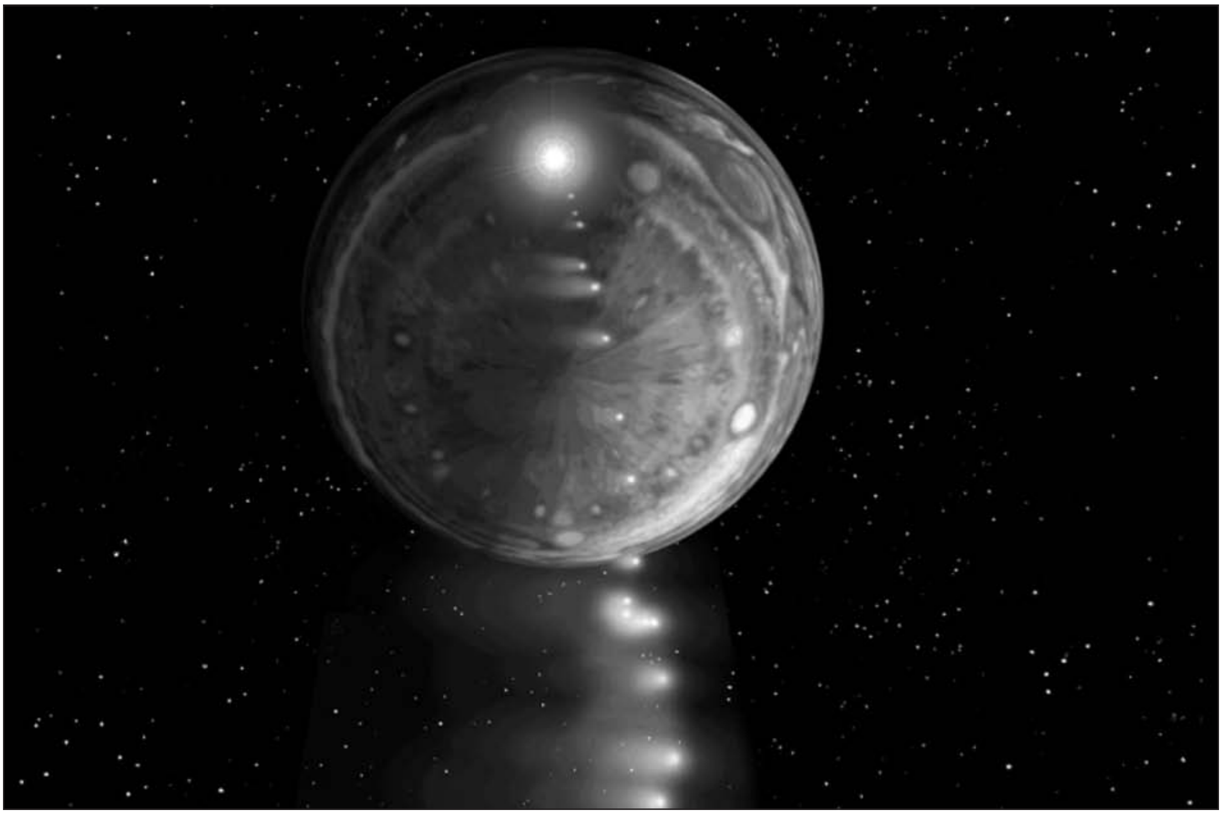
Because astronomers knew when the comet would collide with Jupiter, they had time to organize observatories around the world, including in such remote locations as Antarctica, to make preparations to collect data on this unprecedented event. Because the initial impact would take place on the side of Jupiter hidden from the Earth, several spacecraft would play important roles, particularly in the earliest observations. These spacecraft included Galileo, well on its way to study Jupiter; the Hubble

Space Telescope (HST); Ulysses, which had been studying solar poles; the Roentgen satellite (ROSAT), which had been surveying the sky for X-ray sources; and Voyager 2, which had been exploring the outer planets of the solar system.

As predicted, the comet's first fragments slammed into the Jovian atmosphere on July 16, 1994, at a speed of about 60 kilometers/second, or fast enough to traverse the United States in about a minute. When Jupiter's rotation made the crash site visible to terrestrial observers, thousands of telescopes could see the dark spots that had been created. Fragments of SL 9 continued to collide with Jupiter over the next 5.6 days. Because of the great excitement created by this astronomical event, and because the Earth had been interconnected by various computer networks, images, observations, scientific data, and personal impressions were rapidly transmitted all over the planet. Many others experienced the event through television or through the many stories in magazines and newspapers. By the time the final fragment, W, struck Jupiter on July 22, many millions of Earthlings had shared, in some way or other, this unique interplanetary event.

KNOWLEDGE GAINED

The prodigious wealth of information created by the SL 9 event had important implications for the understanding of comets, the Jovian atmosphere, and the future history of Earth. Astronomers knew, of course, that comets could be destroyed by collisions with the Sun, planets, and satellites, but the data from the fragmentation and collision of SL 9 with Jupiter revealed that its nucleus had been neither a solid body nor a loose agglomeration of materials but something in between. When the pieces hit Jupiter, spectroscopic analysis detected the presence of several elements absent from the Jovian



The remains of Shoemaker-Levy 9 emerge after the comet's impact with Jupiter caused it to break into twenty-one pieces. (JPL/D. Seal, edited by CXC/M. Weiss) atmosphere. These elements, which came from the comet, included such nonmetals as sulfur and silicon and such metals as iron, aluminum, magnesium, and even lithium, hitherto undetected in comets.

As expected, when the high-speed fragments of SL 9 penetrated the Jovian atmosphere, gigantic explosions and massive seismic waves resulted. Fireballs created temperatures in excess of 10,000 kelvins, which rapidly diminished to 2,000 kelvins. Collision-zone temperatures remained elevated for two weeks, but, astonishingly, smaller impact sites had higher temperatures than larger ones. A typical fireball spread from 15 to 100 kilometers in about 40 seconds, and some plumes extended to an altitude of 3,000 kilometers. These explosions also produced waves that sped across the planet at about 450 kilometers/second. These waves, which weakened in about two hours, posed a problem for astronomers. Disagreements developed about where they occurred (in the Jovian stratosphere or troposphere) and how they traveled (guided by a stable layer or generated by interlayer complexities). Just as spectroscopic analysis revealed some surprises about SL 9's chemical composition, so, too, certain elements and compounds were discovered for the first time on Jupiter: for example, diatomic sulfur and carbon disulfide. By contrast, astronomers had expected to find sulfur dioxide, but they did not.

Astronomers also used other parts of the electromagnetic spectrum to gather data on the collision. For example, radio emissions at a specific wavelength (21 centimeters) were indicative of synchrotron radiation, most likely caused by the collision's injecting very-high-speed electrons into the Jovian magnetosphere, which also experienced other changes after the impact. Since both Jupiter and comets were known to have water in their makeup, astronomers were surprised by the very small amounts of water that were detected. Perhaps the comet's fragments lost most of their water before the collision, or perhaps the comet's fragments were destroyed before they reached the planet's water layer.

For many people the highlight of the event was the creation of a series of dark spots that scarred Jupiter's southern hemisphere for several weeks. Similar to the Great Red Spot, if smaller in scale, these dark spots were the most enduring transient features ever seen on the planet, although some historians of science pointed out that, in 1790, Gian Domenico Cassini had reported unusual temporary marks on Jupiter's disk. If these had been due to a cometary collision, then SL 9's crash onto Jupiter would not have been the unique event that many touted it to be.

CONTEXT

Throughout its long history, the Earth has experienced steady and numerous collisions with interplanetary objects. The collision of the Shoemaker-Levy 9 comet with Jupiter provided astronomers with valuable insights into how the collision affected Jupiter and, by analogy, how such comets may have affected other planets, including Earth. Gene Shoemaker estimated that comets had most likely caused about a fifth of the large impact craters on Earth. Linear crater chains have been photographed on Ganymede and Callisto, two of Jupiter's satellites, and these were probably due to cometary collisions. Some scientists have speculated that if SL 9 had collided with Earth instead of Jupiter, a cataclysmic destruction of life would have occurred. According to many scientists, 65 million years ago an asteroid or comet smashed into Central America, creating massive amounts of atmospheric pollutants that helped to bring about the extinction of the dinosaurs and many other forms of life.

In the history of life on Earth other mass extinctions have occurred, and some scientists associate these with periodic comet showers. Various theories have been put forward to explain these periodicities—for example, Nemesis, a companion star of our Sun, may create perturbations in the Oort Cloud that lead to these recurrent invasions of comets into the solar system. Jan Oort was the first to suggest that this large reservoir of icy bodies might be the source of very-long-period comets. However, SL 9's collision with Jupiter revealed something very significant. Because of Jupiter's powerful gravitational field, it attracts many asteroids, comets, and other interplanetary debris, resulting in fewer collisions of these objects with the inner planets, especially Earth. Some have even called Jupiter a "cosmic vacuum cleaner." On the other hand, estimates indicate that small comets collide with Jupiter about once a century, and comets comparable in size to SL 9 hit it about once per millennium. Comet Shoemaker-Levy 9 certainly expanded knowledge about the nature and properties of comets, as well as their interactions with other members of the solar system, but astronomers also realize that they need to learn much more before they will be able to make reliable predictions about some future comet's possibly devastating collision with Earth.

Robert J. Paradowski

FURTHER READING

Fernández, Julio Angel. *Comets: Nature, Dynamics, Origin, and Their Cosmological Relevance*. Dordrecht, Netherlands: Springer, 2005. Using advanced mathematics and celestial mechanics, the author analyzes the history, structure, and behavior of comets, including SL 9. Includes 483 references and an index.

Gehrels, Tom. "Collisions with Comets and Asteroids." *Scientific American* 274, no. 3 (March, 1996): 54-59. A discussion of the likelihood of cometary and asteroid impacts on Earth, designed for the scientifically inclined general audience.

Levy, David H. *Impact Jupiter: The Crash of Comet Shoemaker-Levy 9*. New York: Basic Books, 2003. Called the definitive memoir of SL 9, this book by the comet's codiscoverer provides for the general reader a lively account of the comet, its collision with Jupiter, and the knowledge gained from this event. Illustrated with color and black-and-white photographs. Chapter references and an index.

Noll, Keith S., Harold A. Weaver, and Paul D. Feldman, eds. *The Collision of Comet Shoemaker-Levy 9 and Jupiter*. Cambridge, England: Cambridge University Press, 1996. Contains fifteen reviews by experts involved with the SL 9 event, as well as many references to the primary literature. Index.

Shoemaker, Eugene M., P. R. Weissman, and C. S. Shoemaker. "The Flux of Periodic Comets Near Earth." In *Hazards Due to Comets and Asteroids*, edited by Tom Gehrels. Tucson: University of Arizona Press, 1994. This article explores the probability of a comet like SL 9 colliding with Earth.

Spencer, John R., and Jacqueline Mitton, eds. *The Great Comet Crash: The Collision of Comet Shoemaker-Levy 9 and Jupiter*. Cambridge, England: Cambridge University Press, 1995. A collection of articles by various scientists, including the Shoemakers. Analyzes, from various perspectives, the discovery, tracking, and crash of this comet into Jupiter, as well as discussions of the many discoveries this event stimulated. In the final selections, astronomers explore the possible effects of a comet like SL 9 colliding with Earth.

See also: Callisto; Ceres; Comet Halley; Comets; Dwarf Planets; Eris and Dysnomia; Kuiper Belt; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Pluto and Charon.

Comets

Category: Small Bodies

A comet is a minor body composed mostly of frozen ices typically embedded with solids. Comets revolve about the Sun in highly elliptical orbits. The Oort Cloud is a vast cloud of cometary bodies that extends billions of kilometers out from the Sun.

OVERVIEW

Comets are familiar to nearly everyone as majestic, starlike objects with long tails stretching across a wide band of the sky. The most famous comet, Halley's comet, makes its periodic return to the night skies every seventy-five years. The word "comet" is derived from a Greek word meaning "long-haired." Comets were greatly feared before the twentieth century as

A comet streaks across the night sky. (NASA)

bad omens. Since then, they have been identified and cataloged as objects that come into the inner solar system from deep space. Most of them occupy orbits that carry them far away from the Sun. Many comets make only a single approach to the Sun and then never return again, while others exist in stable, but highly elliptical, orbits that allow them to return after an extended period of time.

One of the first theories advanced to explain the makeup of comets was proposed by astronomer Fred L. Whipple. Whipple suggested that comets were dirty snowballs, essentially bodies of water ice that incorporate dust and perhaps volatives other than water. This remained the primary theory through the first four decades of the space age. Only when spacecraft began visiting comets could it be put to the test.

Until the space age, comets were studied only in visible light through optical telescopic images. The first comet to be studied using Earth-orbital instruments, which permitted observations in the ultraviolet as well as the visible, was the much-heralded Comet Kohoutek late in 1973 and early in 1974. Comet Kohoutek turned out to be something of a disappointment visually from Earth, but images and data collected by the orbiting Skylab 4 astronauts advanced the understanding of comets.

In 1986, the European space probe Giotto passed about 600 kilometers from Halley's comet as the comet made its close approach to the Sun. The probe verified existing theories that comets are made up of ices covered by black dust or soil. In other words, the spacecraft confirmed the dirty snowball model, at least for this comet. Using data taken by the spacecraft, scientists determined that the dust is composed of carbon, hydrogen, oxygen, and nitrogen. Other metals have also been discovered in comets, such as iron, calcium, nickel, potassium, copper, and silicon. Halley's comet was one of the darkest objects ever seen in the solar system; it has virtually no albedo. Only one other major body in the solar system, Saturn's satellite Iapetus, is known to be this low in albedo.

As a comet approaches the Sun, it absorbs solar radiation and becomes warmer. The main body of the comet is called the nucleus. As the nucleus warms, ices beneath the comet's soil evaporate. Because the comet has no atmosphere, evaporated substances, also called volatiles, escape into the vacuum of space. This gaseous envelope that surrounds the comet is called the coma. As the coma grows, it forms a plume of vapor that carries away some of the comet's surface dust as well. This mixture of evaporated volatiles and dust is carried away from the comet by the solar wind, is ionized by high-energy particles, and creates the spectacular tail of the comet. The comet's tail, glowing in the solar wind, can stream behind the comet for millions of kilometers. Cometary nuclei consist mostly of volatile ices and dust. That ice is nearly all water ice, but there is also evidence of ices composed of carbon dioxide and methane. More elementary compounds of nitrogen, oxygen, and carbon monoxide may exist as volatile ices.

Comets are typically small bodies. Halley's comet is an irregular potato-shaped object, 14 by 17 kilometers. In fact, some noted that images of Halley's comet captured during the Giotto mission suggested that the famous comet resembled the cartoon character Felix the Cat. The largest known comet is Chiron, which is estimated to be approximately 200 kilometers in diameter. Comets are thought to have formed as the solar system evolved. Comets were accreted out of material at the outer edge of the solar nebula that ultimately condensed to become the Sun and planets. Because cometary material was fashioned at the outer edge of the solar system, the Sun did not evaporate comets' volatiles. At the same time, the giant planets of the solar system formed at what would become the outer orbits of the solar system. These massive planets encountered the newly formed comets, and the comets that were not engulfed by the giant planets were, over the first billion years, ejected into interstellar space by the planets' massive gravitational fields. Not all comets met that fate, however. Some were gently nudged into stable orbits closer to the Sun. Others were flung into the inner solar system, eventually impacting the inner planets. There are strong reasons to believe that Earth's oceans came from cometary ices delivered to the planet during the early era of bombardment, but that is not universally accepted.

What remained after billions of years of planetary encounters was an extraordinarily large cloud of comets extending outward from orbits beyond Pluto in all directions. A virtual spherically shaped cloud of comets surrounds the Sun at a distance from 1,000 to 100,000 astronomical units (AU). This cloud, which may contain as many as two trillion comets of all shapes and sizes, is called the Oort Cloud. It is named for the Dutch astronomer Jan Hendrik Oort, who first proposed its existence in 1950. The spherically shaped Oort Cloud is not the only source of comets in the solar system. There is a disk-shaped source of comets that extends from about 35 to 40 AU out from the Sun to about 1,000 AU. This source, the Kuiper Belt, was named for the astronomer Gerard Peter Kuiper, who theorized its possible existence in 1951. The disk-shaped Kuiper Belt blends with the spherical Oort Cloud at about 1,000 AU.

The Oort Cloud is the source for long-period comets, with orbital periods of greater than two hundred years. The Kuiper Belt is most likely the primary source for short-period comets, with orbital periods of less than two hundred years, such as Halley's comet. Comets have definite life spans, unlike planets. Each time a comet streaks in toward the Sun, volatile gases stream off the comet and form a beautiful cometary tail, while also depleting the comet's total mass. The comet melts away with each pass toward the Sun. When Halley's comet streamed past the Sun in 1986, the Giotto spacecraft measured a loss of 40 tons of mass per second from the comet. If the supply of comets were not steadily replenished from deep space, they would have all been lost long ago.

The Sun is one among billions of stars in the Milky Way galaxy. In the relatively nearby region of the galaxy, there are hundreds of local stars, which are all revolving around the galactic center and are moving relative to one another. Because stars are so far apart on the average, the chance of one star colliding with another is quite low. However, the possibility of a local star passing near to or through the Oort

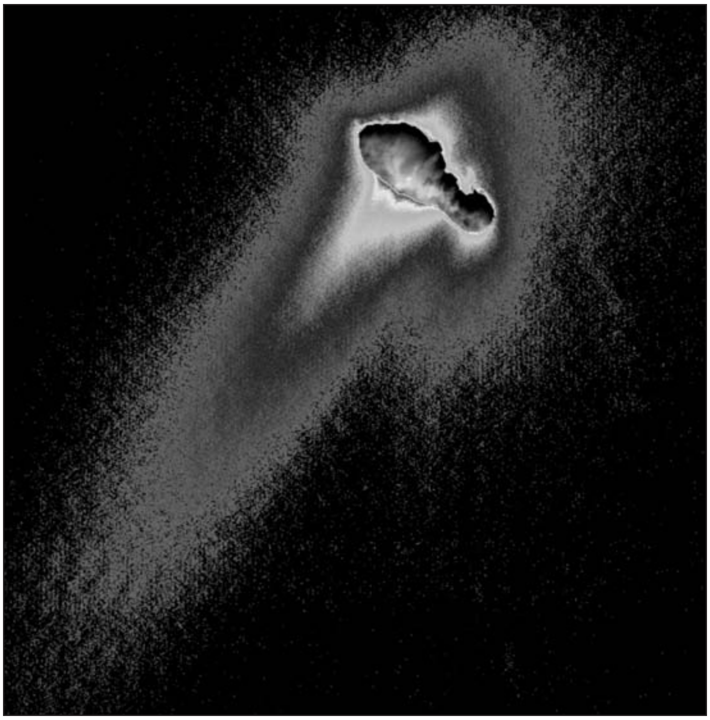
Cloud (which extends up to 100,000 AU away from the Sun) is very high over millions of years. It is estimated that since the solar system formed, about five thousand stars have passed within 100,000 AU of the Sun. If an object as massive as another star passed close to the Oort Cloud, it could easily cause enough gravitational perturbations to direct comets in toward the Sun.

Since the Oort Cloud is spherical, long-period comets can appear to approach the Sun from any point in space. Short-period comets, originating from the Kuiper Belt, always appear to emanate from a band along the ecliptic plane (the plane that contains the planetary orbits). After careful study of where comets actually originate, an analysis was made of their orbits. It has been discovered that there are areas of the sky that are richer in comets than others, and other areas that appear to be practically devoid of comets. Four different theories have been advanced to explain the source of these newly appearing comets. The first theory postulates that the passage of stars in or near the Oort Cloud may so affect the gravitational balance of comets that they are sent falling in toward the Sun. The second theory involves brown dwarfs, which are massive objects—about thirty times the mass of Jupiter—that are not quite planets and not quite stars. They do not have enough mass to create the conditions for thermonuclear ignition at their core. They predominantly radiate infrared energy, and they cannot be readily seen from Earth's surface. Current estimates approximate the number of brown dwarfs near the Sun to be sixty times greater than that of ordinary stars. A brown dwarf should pass through the Oort Cloud every 7 million years. Such an object would travel very slowly with respect to the Sun and would gravitationally release large swarms of comets into the solar system. These two stellar mechanisms, the action of either a passing star or a brown dwarf, are estimated to have been the source of about one-third of the observed comets.

Facts About Selected Comets

Name (no)	*Period (yrs)*	*Perihelion Date*	*Perihelion Distance*	*Distance from Sun (AU)*
Borrelly (19P)	6.86	2001-09-14	1.358	3.59
Chiron (95P)	50.7	1996-02-14	8.460	13.7
Crommelin (27P)	27.89	1984-09-01	0.743	9.20
d'Arrest (6P)	6.51	2008-08-01	1.346	3.49
Encke (2P)	3.30	2003-12-28	0.340	2.21
Giacobini-Zinner (21P)	6.52	1998-11-21	0.996	3.52
Grigg-Skjellerup (26P)	5.09	1992-07-22	0.989	2.96
Hale-Bopp	4,000	1997-03-31	0.914	250
Halley (1P)	76.1	1986-02-09	0.587	17.94
Honda-Mrkos-Pajdusakova (45P)	5.29	1995-12-25	0.581	3.02
Hyakutak	~40,000	1996-05-01	0.230	~1,165
Kohoutek (75P)	6.24	1973-12-28	1.571	3.4
Schwassmann-Wachmann 3 (73P)	5.35	2006-06-02	0.933	3.06
Tempel 1 (9P)	5.51	2005-07-07	1.497	3.12
Tempel-Tuttle (55P)	32.92	1998-02-28	0.982	10.33
West-Kohoutek-Ikemura (76P)	6.46	2000-06-01	1.596	3.45
Wild 2 (81P)	6.39	2003-09-25	1.583	3.44
Wilson-Harrington (107P)	4.30	2001-03-24	1.000	2.64
Wirtanen (46P)	5.46	2013-10-21	1.063	3.12

Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.



A composite of images captured by Deep Space 1, about 4,800 kilometers from Comet Borrelly, shows the comet's nucleus during the spacecraft's September, 2001, flyby. (NASA/JPL)

According to the third theory, huge molecular clouds in interstellar space (much more massive than a single star) may pass at very large distances (tens of light-years) and may cause a release of comets through gentle perturbations of their orbits. The final theory for the source of newly appearing comets is galactic tidal action. Each galaxy has a gravitational field, which causes an attraction toward the midplane of the galaxy of all bodies (comets and stars). As these bodies orbit the galaxy, they are gravitationally influenced by one another. The galactic tide is the difference between the galactic forces acting on the Sun and the comet. Because the force of the galactic tide is very specific with respect to direction, it cannot act toward the poles of the Sun or toward the equator. Observations of cometary tracks confirm that comets from deep space do not seem to approach the Sun from these segments of the celestial sphere. This mechanism appears to explain the approach of the majority of all long-period comets entering the solar system from the Oort Cloud.

In the aftermath of the American decision to be the only spacefaring nation not to dispatch a spacecraft to investigate Halley's comet on its most recent appearance in the inner solar system, the National Aeronautics and Space Administration (NASA) proposed the Comet Rendezvous and Asteroid Flyby (CRAF). CRAF was a sister ship to the Cassini spacecraft. Because of budget cuts, NASA was able to save the Cassini mission, but CRAF was canceled. If it had been adopted, CRAF would have rendezvoused with Comet Kopff and remained in its vicinity for thirty-two months to observe variations in that comet's activity during different portions of its orbit. CRAF would also have dropped penetrometers into the comet to ascertain information about internal structure and make chemical analyses of surface materials.

Aspects of the ambitious CRAF concept were recycled into cheaper comet missions such as Deep Space 1, Stardust, and Deep Impact. Also, the European Space Agency (ESA) developed the Rosetta spacecraft to visit a comet.

Launched on October 24, 1998, the Deep Space 1 spacecraft began tests of an ion propulsion system and autonomous navigation system. Deep Space 1's targets were the asteroid Braille and Comet Borrelly. Flying by the comet at a relatively close distance, Deep Space 1 captured images of a comet's nucleus that had higher resolutions than any of those captured by the probes that had visited Halley's comet. Comet Borrelly was shaped much like a bowling pin and displayed emission jets not distributed uniformly across its irregular nucleus. Deep Space 1 was not outfitted with debris shields, but it survived the close encounter nevertheless. In time Deep Space 1 ran out of propellant, but the mission showed that ion propulsion could be used on a spacecraft designed to visit multiple targets such as comets.

Stardust was launched on Feburary 7, 1999, and directed toward Comet Wild 2, where it opened up special sample collectors incorporating aerogel to capture both interplanetary and cometary dust. After the spacecraft's encounter with the comet on January 2, 2004, its sample collectors were sealed for a two-year journey back to Earth. On January 15, 2006, Stardust's sample collection unit safely reentered Earth's atmosphere and was recovered intact in Utah.

In June, 2008, researchers studying comet 26P/Grigg-Skjellerup material collected by the Stardust spacecraft announced that they had discovered new mineral grains. This mineral, named brownleeite after Donald Brownlee of the University of Washington, was a variety of manganese silicide not previously predicted by models of comets or the condensation of material from the early proto-Sun's nebula.

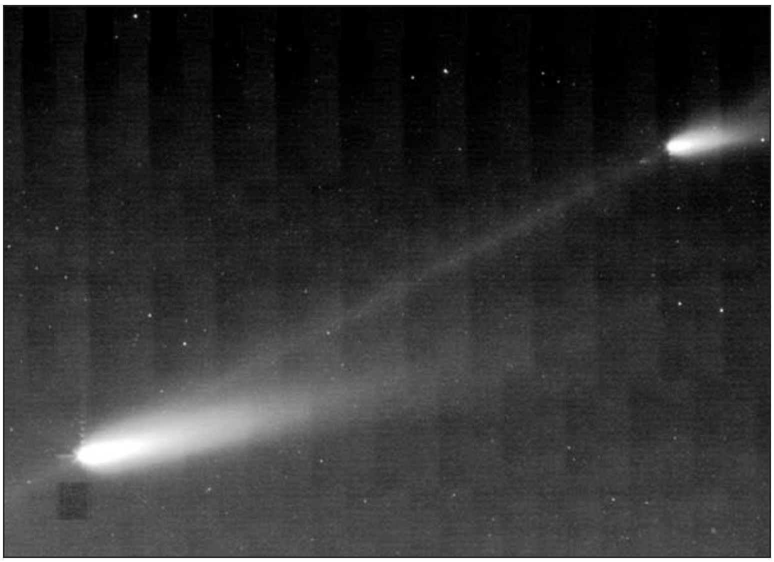
Deep Impact was designed as its name clearly suggests to fly to a comet and strike it, excavating material from deep below the surface. The spacecraft launched on January 12, 2005, and its onboard navigation steered the spacecraft toward the comet Tempel 1, released an impactor payload made largely of copper, and then veered out of the way to observe the resulting impact of the payload on the comet. The impactor was composed of copper, since that was an element not expected to be found naturally within the comet. The impact was observed by Deep Impact itself, as well as by the Hubble Space Telescope, the Chandra X-Ray Orbservatory, the Spitzer Space Telescope, the Swift spacecraft, and ESA's XMM-Newton observatory and Rosetta spacecraft. This coordinated effort permitted time-evolution studies of the plume and debris cloud created by the high-speed impact of the copper payload on Tempel 1.

Rosetta was launched on March 2, 2004. This was the European Space Agency's second attempt at a comet study and incorporated both a flyby craft and a lander named Philae. The mission was designed to rendezvous with the comet 67P/Churyumov-Gerasimenko in May, 2014, orbit it for many months while mapping the surface, and observe changes in the comet's activity as its distance to the Sun changed. Then the lander was scheduled to touch down on the comet on or about November, 2014, where it would secure itself to the surface in the comet's weak gravity field and then begin studies of chemical composition and physical characteristics of the comet's surface. The Rosetta mission was planned to continue through December, 2015.

APPLICATIONS

The study of comets requires detailed knowledge of the composition of the outer regions of the solar system and the space between the last planet and 100,000 AU outward from the Sun. Comet studies also seek to understand complex gravitational interactions between bodies separated by wide distances and even gravitational interactions between tiny comets and the entire galaxy. Astronomers who study comets want to learn more about their makeup, their behavior when approaching the Sun, and the makeup and evolution of the early solar system.

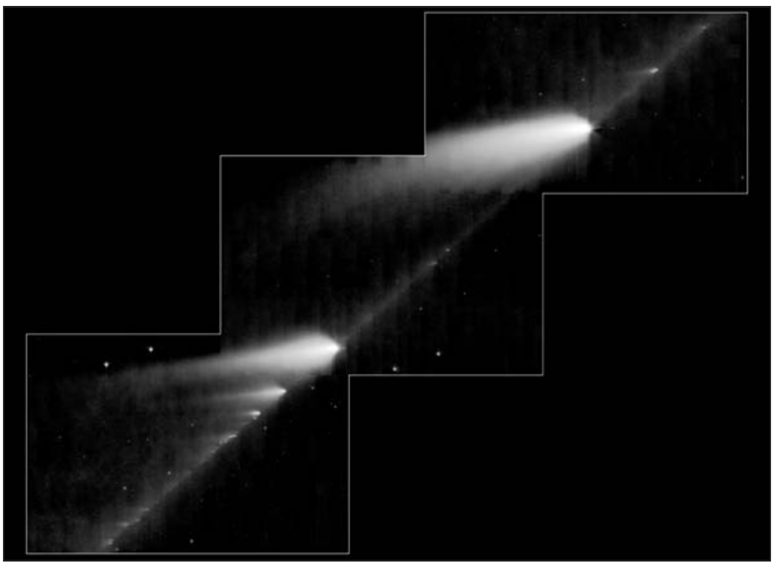
New comets approaching the Sun for the first time have been held in deep freeze within the Oort Cloud and are thought to be composed of primordial material of the newly forming solar



The Spitzer Space Telescope captured this infrared image of the breakup of Comet 73P/Schwassman-Wachmann 3, which began to split into pieces in 1995. (NASA/JPL-Caltech) system. They have been tied up in the Oort Cloud for billions of years at temperatures barely above absolute zero. As they approach the Sun, their internal gases begin to stream away. Detailed study of an approaching comet's outgassing can inform planetary scientists about the composition of the early solar system. Comets and their approach have also hinted at the existence of the elusive brown dwarfs, thought to be one of the most common bodies of interstellar space. Because they are so dim, they are all but invisible from Earth. On the other hand, because brown dwarfs are thought to be so plentiful, the study of comets and their orbits may give the first real clues to the former's reality and abundance. The first serious studies of brown dwarfs came from observations made by the Spitzer Space Telescope, the final member of NASA's Great Observatory program. Spitzer detected brown dwarfs from their infrared emissions.

In the early 1980's the existence of galactic tidal action was merely speculation. Since then, careful study of comet orbits and their approaches to the inner solar system has favorably supported the theory of galactic tides. In the close approach of Halley's comet by robotic spacecraft in 1986, a wealth of information was recovered on the shape, behavior, and composition of comets. The existence of the Oort Cloud and the concept of gravitational interactions by passing objects in space have led to the theory of periodic comet showers. Such comet showers, separated by periods of tens of millions of years, may be responsible for certain mass extinctions on Earth. These extinctions might be the result of a shower of comets originating from within the Oort Cloud, sent on their close approach to the Sun by the close passage of a star or brown dwarf to or through the Oort Cloud.

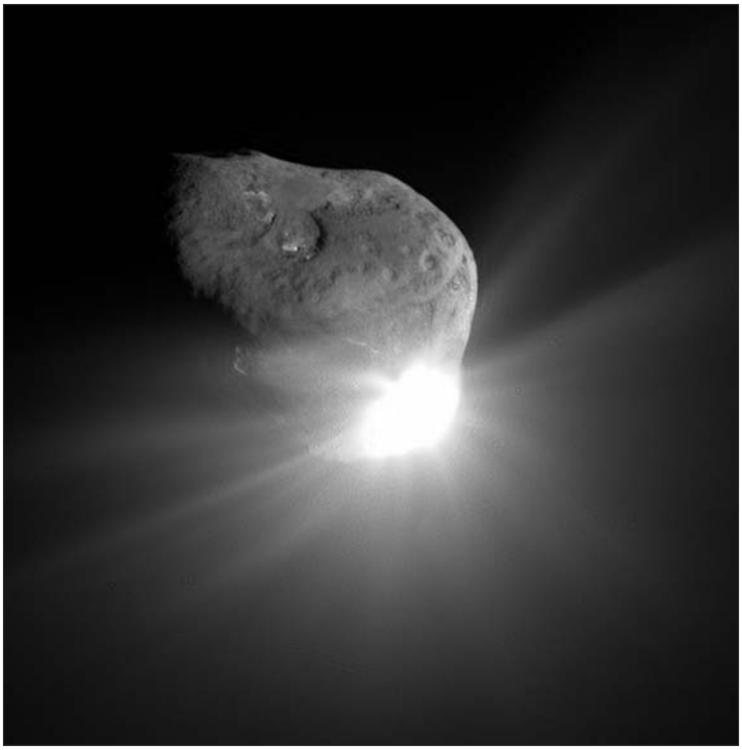
Samples of Comet Wild 2 were treated to many of the contamination safeguards used with the Apollo lunar rocks. Analyses of Stardust's captured comet material revealed some surprises. Tracks in the aerogel suggested solid materials larger than interstellar dust grains. Silicate crystals and other mineral crystals were found which required more than just mild heating, as would have been the case if the comet was largely composed of interstellar dust grains. This suggested that the theory of comet formation may need alteration. Inclusions of vanadium nitride, titanium, molybdenum, osmium, ruthenium, and tungsten were found, components that would have required high heating. Samples also contained organic materials more primitive than found in asteroidal material, compounds such as polycyclic aromatic hydrocarbons.



Debris from comet 73P/Schwassman-Wachmann 3, taken in the infrared by Spitzer. The debris passes near Earth every year and is expected to cause a significant meteor-shower display in the year 2022. (NASA/JPL-Caltech)

The collision of Deep Impact's copper payload was equivalent to five tons of TNT. Comet Tempel 1 increased in brightness sixfold as a result of the event. As much as between 10 and 25 million kilograms of comet material was ejected as a crater formed as a result of the impact. Tempel 1 material was much finer than had been expected, being more akin to talcum powder than a sandy grain. Data ruled out a loose aggregate or highly porous model of the comet's structure. Indeed, rather unlike a dirty snowball as proposed by Whipple, Comet Tempel 1 was more like an icy dirt ball. Seen in the ejected materials in addition to volatiles were clays, carbonates, sodium, and crystalline silicates. After this encounter the flyby portion of the Deep Impact spacecraft was redirected to a planned encounter with Comet Hartley 2 in 2010.

Rosetta holds the potential for in situ analyses of comet material if the Philae lander successfully touches down safely, and can perform its chemical and physical tests. This information will greatly assist in determining how comets form and how their comas and tails develop as they travel into the inner solar system.



Comet Tempel 1 in an image taken in 1998 by the Deep Impact flyby craft. (NASA/JPL-Caltech)

CONTEXT

Humankind has always looked to the heavens in awe and wonder, and sometimes in fear. Perhaps no other astronomical phenomenon except a total solar eclipse has historically evoked as much fear as comets. When the specter of fear is removed, however, they emerge as strikingly beautiful objects in the sky. It was once believed that if Earth passed through the tail of a comet, its inhabitants would die; this theory has been discredited. Comets are messengers from a time long past. Most are chunks of dirty ice, locked away in the Oort Cloud for billions of years.

Comets have been used to judge vast distances, evaluate the early composition of the solar system, and even test the idea that the gravity of the entire Galaxy can make a difference to the smallest of objects in space. Comets have been used as yardsticks to evaluate what may be the most common type of star in the galaxy, the brown dwarf—which ironically is one that is difficult to observe, even in the infrared. Comets have been called dirty snowballs. Halley's comet was so black that it was the darkest object ever seen in space. Yet, from these dirty specks of ice, planetary scientists have witnessed some of the most spectacular light shows. Ultimately, comets may also generate clues to some of the most fundamental secrets about the solar system and planets. From these tiny messengers, planetary scientists may unlock and examine pristine elements from the formation of the solar system.

Debris from comets provides the material that Earth passes through when annual meteor showers occur. For example, the Orionid meteor shower is leftover material from Halley's comet, the Leonids meteor shower is associated with Comet Tempel-Tuttle, and the Perseid meteor shower is material from Comet Swift-Tuttle.

Historically comets have come full circle from being seen as omens in the heavens to be feared, to celestial objects evoking a sense of wonder, and to again being objects that should be feared if they come too close and perhaps even impact Earth. Comets represent a more troubling threat than do asteroids, as comets are usually discovered only when they come in past the orbit of Jupiter. As such there is insufficient time to mount any mitigating effort if a new comet is determined to make a close pass or actually impact the Earth. Have comets impacted the Earth in the past? The answer is believed to be almost certainly yes, and indeed many believe that the majority of Earth's water came from comets encountering the early Earth. The Siberian Tunguska event of 1908, itself a curiosity that has been explained by some (without any legitimate supporting evidence) as a nuclear explosion or even the impact of an unidentified flying object, is now believed to have been the result of a comet or asteroid impact, most likely an air burst explosion of the body. Although this theory is not yet confirmed, it points out the potential for devastation that an impacting comet represents.

A comet collision was observed in 1994 when the nearly two dozen pieces of the shattered Comet Shoemaker-Levy 9 smacked into Jupiter's upper atmosphere. These pieces created temporary changes in the gas giant's appearance, many of which were the size of the Earth, indicating that a tremendous amount of energy was involved in this series of collisions. The impacts were recorded by the Hubble Space Telescope and Galileo spacecraft; the incredible magnitude of the disruption came as a surprise to the scientific community.

It was believed since roughly 1950 that gravitational disruption of the Sun's Oort Cloud by a close passage of another star was responsible for swarms of comets heading into the inner solar system, resulting in bombardment of the planets. However, in late 2008 Hans Rickman of Sweden's Uppsala Astronomical Observatory reported in *Celestial Mechanics and Dynamical Astronomy* the results of an updated computer simulation of the Oort Cloud investigated by his research group. If correct, their model indicates that sporadic stellar encounters, while indeed important in generating fresh comets that head toward the inner solar system, is not the only mechanism for sending comets toward the planets. This model accounted for galactic gravitational tidal influences on the Oort Cloud and found that the threat from comets may be more constant than previously believed. If correct, the model reinforces the need to monitor the skies for incoming comets that might be headed our way.

Dennis Chamberland and David G. Fisher

FURTHER READING

Arny, Thomas T. *Explorations: An Introduction to Astronomy*. 3d ed. New York: McGraw-Hill, 2003. A general astronomy text for the nonscience reader. Includes an interactive CD-ROM and is updated with a Web site.

Benningfield, Damond. "Where Do Comets Come From?" *Astronomy* 18 (September, 1990): 28-36. A fine summary of comets, superbly illustrated and written for the general public. Addresses the question of the Oort Cloud and Kuiper Belt in detailed, scaled illustrations. Discusses possible linkage to the extinction of the dinosaurs and the latest satellite discoveries.

Brandt, John C., and Robert D. Chapman. *Introduction to Comets*. New York: Cambridge University Press, 2004. A text suitable for a planetary science course, this comprehensive work covers our knowledge of comets from early observations to telescopic investigations through spacecraft encounters of these mysterious and alluring bodies.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Well-explained material on asteroids and comets. Takes a comparative planetology approach.

Levy, David H. *The Quest for Comets: An Explosive Trail of Beauty and Danger*. New York: Plenum Press, 1994. Written by one of the codiscoverers of Comet Shoemaker-Levy 9, this book is for the general reader. It highlights the author's comet discovery program and the comet catastrophe theory.

Newburn, R. L., M. Neugebauer, and Jurgen H. Rahe, eds. *Comets in the Post-Halley Era*. New York: Springer, 2007. A collection of fifty papers compiled into a volume in Springer's Astrophysics and Space Science Library. Covers observational techniques, origin of comets, evolution of comets, and spacecraft data. Special attention is given to Comet Halley and other comets encountered by spacecraft.

Russell, Christopher T. *Deep Impact Mission: Looking Beneath the Surface of a Cometary Nucleus*. New York: Springer, 2005. A complete description of the Deep Impact mission to excavate and analyze material from Comet 9P/Tempel 1.

Sagan, Carl, and Ann Druyan. *Comet*. New York: Random House, 1985. This coffee-table book is a classic work of art, written by the most popular astronomer in the United States. Filled with beautiful color and historical black-and-white photographs and illustrations, it was the basis for the popular movie by the same name. For general audiences.

Schaaf, Fred. *Comet of the Century: From Halley to Hale-Bopp*. New York: Copernicus Books, 1997. Comet Hale-Bopp was a popular comet to observe, one that provided a better show than the most recent appearance of Halley's comet. For a general audience.

Thomas, Paul J., Roland D. Hicks, Christopher F. Chyba, and Christopher P. McKay. *Comets and the Origin and Evolution of Life*. 2d ed. New York: Springer, 2006. A collection of chapters written by experts in the field. This update of the first edition covers new understandings of Halley's comet and more recent spacecraft data. Provides insights into organic compounds found in comets, protostars, and interstellar clouds.

Verschuur, Gerrit L. *Impact! The Threat of Comets and Asteroids*. New York: Oxford University Press, 1997. Verschuur explains the change in thinking from uniformitarianism to catastrophism. Identifies the Chicxulub Crater with an impact event that led to the extinction of the dinosaurs. Warns of the potential for devastation that a comet impact on Earth would cause.

See also: Ceres; Comet Halley; Comet Shoemaker-Levy 9; Dwarf Planets; Earth's Oceans; Eris and Dysnomia; Extraterrestrial Life in the Solar System; Infrared Astronomy; Kuiper Belt; Lunar Craters; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Planetary Classifications; Planetary Formation; Pluto and Charon; Solar System: Element Distribution; Solar System: Origins; Solar Wind; Ultraviolet Astronomy.

Coordinate Systems

Category: Scientific Methods

There are several astronomical coordinate systems that are in common usage. In each system, the position of an object in the sky, or on the celestial sphere, is specified by two angles, similar to latitude and longitude on Earth.

OVERVIEW

An astronomical coordinate system is a way for locating an object in the sky, or on the celestial sphere, using two angles. (The celestial sphere is an imaginary sphere of large size surrounding the Earth, representing the sky as seen from Earth.) There are several astronomical coordinate systems in common usage, and each system is based on a reference plane and a reference direction in that plane. In each system, the intersection of the reference plane and the celestial sphere is a great circle on the celestial sphere defining the "equator" of the coordinate system. The two "poles" of the system are the two points on the celestial sphere each 90º from the system's equator. Great circles passing through these poles intersect the equator of the system at right angles. One of the two angular coordinates of each coordinate system is measured from the equator of the system to the object along the great circle passing through it and the poles. Angles on one side of the equator are considered positive; those on the opposite side are negative. The other angular coordinate is measured along the equator from the reference direction to the intersection of the equator, with the great circle passing through the object and the poles. (For comparison, in the system of latitude and longitude on Earth, the reference plane is the Earth's equatorial plane, and the reference direction is the intersection of the prime meridian passing through Greenwich, England, with the equator. Latitude is measured as the angle north or south of the equator, and longitude is measured as the angle east or west of the prime meridian).

Four astronomical coordinate systems are commonly used: alt-azimuth (or horizon), equatorial, ecliptic, and galactic.

The alt-azimuth, or horizon, system has as its reference plane the plane of the horizon, which is a great circle on the celestial sphere 90° from the zenith (the point directly over the observer). Its reference direction is the north point (the point on the horizon due north). Its latitude-like coordinate, called altitude (h), is the angle above or below the horizon (positive above and negative below). Altitude ranges from +90° at the zenith to –90° at the nadir (the point directly underneath the observer). The longitude-like coordinate, called azimuth (A), is the angle measured to the east along the horizon from the north point. Azimuth varies from 0° due north, to 90° due east, to 180° due south, to 270° due west, and to 360° (equivalent to 0°) as north is approached from the west. This system is convenient for giving approximate directions to objects in the sky at any one moment from any one location on Earth, but the altitude and azimuth for any object are different as seen from different locations on Earth, and they constantly change as objects appear to move across the sky due to the Earth's rotation.

In the equatorial system, the reference plane is the plane of the Earth's equator; its intersection with the celestial sphere defines the celestial equator. The extension of the Earth's rotational axis through the north and south terrestrial (or geographic) poles intersects the celestial sphere at the north and south celestial poles. The reference direction is given by the Sun's apparent position at the moment of the vernal equinox, when the Sun is directly above some point on the Earth's equator in March. The latitude-like coordinate, called declination, is measured as the angle north or south of the celestial equator (positive to the north and negative to the south). Declination ranges from +90° at the north celestial pole (the point on the celestial sphere directly over the Earth's north pole) to –90° at the south celestial pole (the point on the celestial sphere directly over the Earth's south pole). The longitude-like coordinate, called right ascension, is measured to the east along the celestial equator from the vernal equinox point. Instead of specifying this angle in degrees, right ascension is given traditionally in hours, minutes, and seconds. Because of the Earth' s rotation on its axis, the sky appears to turn through 360° in approximately 24 hours: exactly 24 hours of sidereal time, but 23 hours, 56 minutes, and 4 seconds of mean solar time (the difference in these two time systems is due to the Earth's revolution around the Sun). A full circle of 360° around the celestial equator is defined as 24 hours of right ascension. Thus 1 hour of right ascension corresponds to 15° of arc, 1 minute of right ascension corresponds to 15 arc minutes, and 1 second of right ascension corresponds to 15 arc seconds.

This equatorial system rotates with the apparent daily motion of the sky, so declination and right ascension of most astronomical objects remain nearly constant for reasonably short times of up to a few years. However, over the long run, declination and right ascension change as a result of two main factors: precession and proper motion. (Note that the Sun, Moon, and planets appear to move through the sky with respect to the stars, so their right ascension and declination noticeably change much more rapidly, over timescales of hours to months.) Precession is a slow change in the direction of the Earth's rotational axis, which traces out a double-cone figure in space over approximately 25,800 years, due to the gravitational pull of the Moon and Sun on the Earth's equatorial bulge. As a result, the positions of the celestial poles and celestial equator do not remain fixed on the celestial sphere, but they shift in a predictable way. Each of the two celestial poles traces out a circle with an angular diameter of 47° on the celestial sphere over a period of 25,800 years. Because of precession, the equinoxes shift westward relative to the stars so that gradually the seasonal constellations change. The vernal equinox (occurring in March), now is located in Pisces, but 2,000 years ago it was in Aries, and in about 600 years it will shift into Aquarius. Polaris currently is the north pole star, but at the time of the building of the Great Pyramid in Egypt, about 5,000 years ago, the star Thuban in Draco was the north pole star. In about 12,000 years, the star Vega in Lyra will approximately mark the north celestial pole. Because of precession, catalogs listing the declination and right ascension of stars, nebulae, galaxies, and other objects must specify the epoch (year) for which the coordinates are rigorously correct. Corrections must be calculated using standard precessional formulas to convert the listed coordinates to other years.

The declination and right ascension of individual stars also change due to their proper motion, which is the change in direction to a star as seen from Earth due to the star's actual motion through space relative to our solar system. For most stars, the proper motion is small enough so that its practical effect on declination and right ascension can be ignored, but some stars have proper motions of several arc seconds per year (the largest proper motion known is that of Barnards star, 10.4 arc seconds per year), so over timescales of decades the declination and right ascension of these stars can change by arc minutes. (A very long-term effect of proper motions is that over timescales of thousands of years or more, the familiar shapes of constellations slowly change.)

In the ecliptic system, the reference plane is the plane of the Earth's orbit around the Sun. The term "ecliptic" is used both for the Earth's orbital plane and for the great circle that marks its intersection with the celestial sphere. The reference direction is the same as in the equatorial system: the apparent position of the Sun at the moment of the vernal equinox. The term "vernal equinox" is used both for the moment in time in March when the Sun is directly above some point on the Earth's equator and for one of the two intersections of the ecliptic and the celestial equator on the celestial sphere. (The other intersection is called the autumnal equinox and marks the Sun's apparent position when it is directly above some point on the Earth's equator in September.) The latitude-like coordinate, called ecliptic latitude, is measured as the angle north (positive) or south (negative) of the ecliptic; it ranges from +90° to –90°. The longitude-like coordinate, called ecliptic longitude, is measured to the east along the ecliptic from the vernal equinox point; it ranges from 0° to 360°. This coordinate system is especially useful in giving the position of solar-system objects. Because many solar-system objects have orbits around the Sun that are not inclined very greatly to the plane of the Earth's orbit around the Sun, such solar-system objects will be seen near the ecliptic and thus have small ecliptic latitudes.

The galactic coordinate system has a reference plane that is the mean plane of our galaxy, the Milky Way (defined primarily by 21-centimeter, 1,400-megahertz radio observations of neutral hydrogen, which is concentrated in the galactic plane), and a reference direction that points to the galactic center. The latitude-like coordinate, called galactic latitude, is measured as the angle north (positive) or south (negative) of the galactic equator. The longitude-like coordinate, called galactic longitude, is measured along the galactic equator from the direction to the galactic center (as viewed from the poition of our solar system) toward the direction of galactic rotation. It varies from 0° in the direction of the galactic center, to 90° in the direction of general galactic rotation, to 180° in the direction opposite to the galactic center (the galactic "anticenter"), to 270° opposite to the direction of general galactic rotation, and up to 360° as the direction to the galactic center is approached again. This system is useful for indicating the location of objects relative to the Milky Way.

APPLICATIONS

The primary use of any astronomical coordinate system is to specify the location of celestial objects for observation. Most sources list the equatorial coordinates of right ascension and declination, together with the epoch (year) for which the right ascension and declination are rigorously correct. In the past, astronomers had to calculate precessional corrections to adjust the right ascensions and declinations to the current date, and then calculate the angles to set the telescope based on the time of observation. However, now the settings of most major telescopes have been computerized so that the right ascension, declination, and listing epoch can be input directly, and the correct settings will be calculated automatically.

Another use for old catalogs listing right ascension and declination is to determine proper motions. The old coordinates, corrected for precession, are compared to new coordinates. Any differences are due to the proper motions of the objects.

In the case of an object in our solar system (such as a planet, asteroid, or comet), often the elements of its orbit around the Sun (such as semimajor axis, eccentricity, inclination, or time of perihelion passage) will be given. These can be used to calculate the object's ecliptic latitude and longitude for a specific time and date, and in turn these can be converted into its right ascension and declination for that same time and date. A list of these coordinates for a series of dates provides an ephemeris of where to observe the object.

In the study of the structure of the Milky Way, coordinates of galactic latitude and galactic longitude often are most useful for visualizing where objects are located. Of course, galactic coordinates can be converted to equatorial coordinates of right ascension and declination for setting telescopes to observe the objects.

CONTEXT

The earliest references to locations and motions of stars and planets can be traced back about twenty-five hundred years, when Babylonian observer-priests recorded the movement of planets relative to the stars. In the hands of the Greeks, the Babylonian results for periodic and irregular celestial motions became the basis for geometric models trying to explain the structure and motions of the universe. Ptolemy, between 296 and 272 B.C.E., measured positions for a number of stars in terms of their angular distances above or below the celestial equator (today this would be called declination), as well as differences in their angular positions parallel to the celestial equator (today this would be called differences in right ascension). Around 150 B.C.E., the Greek astronomer Hipparchus compiled one of the first systematic star catalogs, giving the coordinates and apparent magnitudes for about 850 stars.

The development of astronomical coordinate systems was important for several practical reasons. The most significant was that it allowed for the construction of an accurate calendar, which was essential for weather prediction and agriculture. The length of the year could be fixed, months and days could be intercalated, and the passing of the solstices and equinoxes could be established. In addition, astronomical coordinate systems led to the development of celestial navigation, enabling commerce to expand to new trade areas and allowing more of the world to be explored.

Earl G. Hoover and Richard R. Erickson

FURTHER READING

Boucher, C., ed. *Earth Rotation and Coordinate Reference Frames*. International Association of Geodesy Symposia 105. New York: Springer, 1990. Discusses techniques for determining Earth's rotation, including laser ranging and very long baseline interferometry (VLBI).

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Includes a section on coordinates.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook, thorough and well written. Includes a section on celestial coordinates.

Kovalevsky, Jean. *Modern Astrometry*. New York: Springer, 1995. An introduction to astrometry. Covers how to find motions, positions, dimensions, and other data about astronomical objects. Also discusses observational techniques and equipment.

Lankford, John, ed. *History of Astronomy: An Encyclopedia*. New York: Garland, 1997. An encyclopedia focused on the history of astronomy, women in astronomy, social aspects, and national contexts.

Maran, Stephen P., ed. *The Astronomy and Astrophysics Encyclopedia*. Foreword by Carl Sagan. New York: Van Nostrand Reinhold, 1992. More than four hundred articles arranged alphabetically. Full of diagrams and photographs. Easy to use and read.

Moche, Dinah L. *Astronomy: A Self-Teaching Guide*. 6th ed. New York: John Wiley & Sons, 2004. A self-instructional text designed so that students with no formal astronomy background can easily learn basic principles and concepts. The material in each chapter is presented in short, numbered sections. The chapter on understanding the starry night is especially recommended for its coverage of the coordinate systems. An excellent book for the upper-level high school and lower-level college students.

Pannekoek, A. *A History of Astronomy*. London: Barnes & Noble Books, 1969. As the title denotes, this is a history and as such is written in a nonscientific style. Of special interest are the early chapters covering the Babylonians, Assyrians, and Chaldean contributions. Recommended for the college-level student and interested general reader.

Pasachoff, Jay M., and Will Tirion. *Field Guide to the Stars and Planets*. 5th ed. Boston: Houghton Mifflin, 1999. An excellent handy reference. Suitable for all high school upper-level students, college students, and hobbyists. Useful tables give star names and coordinates. In addition, one chapter has an easy-to-read discussion on coordinates, time, and calendars.

Rey, H. A. *The Stars: A New Way to See Them*. 1952. Reprint. Boston: Houghton Mifflin, 1988. A simple, clearly written, and charming book. Includes a good introduction to celestial coordinates and time. Written by the author of the "Curious George" children's stories.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses, divided into many short sections on specific topics. Contains sections on coordinate systems. However, being the first edition of a revised text, there are several typographical errors throughout.

See also: Earth System Science; Earth's Rotation; Gravity Measurement; Hertzsprung-Russell Diagram; Infrared Astronomy; Neutrino Astronomy; Optical Astronomy; Radio Astronomy; Telescopes: Ground-Based; Telescopes: Space-Based; Ultraviolet Astronomy; X-Ray and Gamma-Ray Astronomy.

Coronal Holes and Coronal Mass Ejections

Category: The Sun

The corona is the outermost layer of the Sun. It is extremely hot but so tenuous that it is visible only when a solar eclipse blocks the brighter photosphere. Coronal holes are less dense regions of the corona where coronal matter streams outward into interplanetary space. Coronal mass ejections occur when magnetic field lines in the solar corona snap and eject large clumps of solar material into interplanetary space.

OVERVIEW

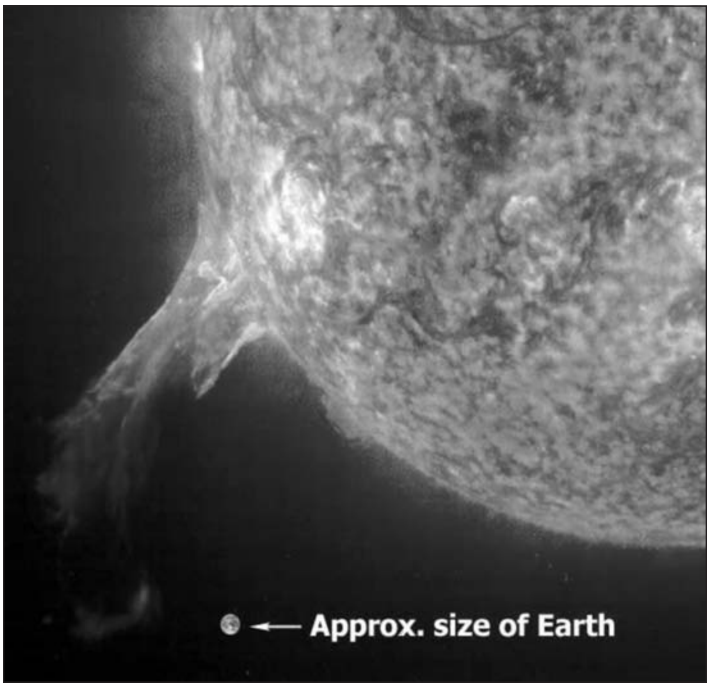
The outermost layer in the Sun's atmosphere is the corona (which means "crown"). Gas in the corona can reach temperatures of a few million kelvins. This gas is very thin, however, with a density on the order of 10^{-12} kilograms/meter^3. Thus, the corona is faint, so faint that it cannot normally be seen because its feeble light is overwhelmed by the much brighter photosphere. The corona must be observed optically either during a total solar eclipse or by using a coronagraph. The latter is a disk, blocking the photosphere, in the focal plane of the telescope.

The chromosphere is the layer of the Sun's atmosphere between the photosphere and corona. The chromosphere is only a few thousand kilometers thick. The temperature of the gas in the chromosphere is slightly higher than that in the photosphere. In the approximately 100-kilometer-thick transitional region between the chromosphere and corona, the temperature rapidly increases from about 6,000 kelvins to a few hundred thousand kelvins.

The Sun is a ball of gas without anything that could be considered a solid surface. The Sun's photosphere, however, is the closest the Sun has to a surface. The photosphere is relatively opaque and blocks our view of the solar interior, so most photographs or observations of the solar disk show the Sun's photosphere. It is also the coolest layer of the Sun. The bottom layer of the photosphere is at a temperature of 5,800 kelvins. The photospheric temperature drops with increasing height to a temperature of about 4,500 kelvins at the top of the photosphere, begins to increase in the chromosphere, and is extremely high in the corona.

Relatively cool stars are reddish in color, while hot stars are bluish. The Sun's corona, at a few million kelvins, is much hotter than most stars, so it emits most of its energy in the extreme ultraviolet (the shortest ultraviolet wavelengths) to X-ray region of the electromagnetic spectrum. The photosphere is not hot enough to emit significant amounts of energy in this spectral region. Fortunately for human beings, Earth's atmosphere blocks most extreme ultraviolet and X radiation, so astronomers study the Sun at these wavelengths from satellites.

Extreme ultraviolet and X-ray pictures of the Sun show a bright corona and dark photosphere, which is the reverse of optical pictures showing a bright photosphere and much fainter corona. Solar astronomers therefore study the solar corona using extreme ultraviolet or X-ray images. At these wavelengths the corona shows structures that are not visible at optical wavelengths.



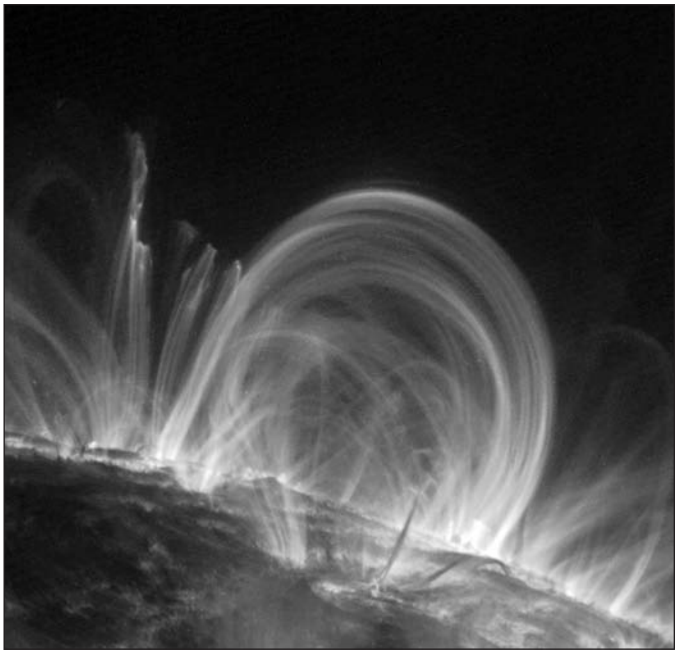
This huge CME was imaged in 2003 by the SOHO spacecraft. (NASA/European Space Agency)

One common structure in the X-ray corona is the coronal hole. On X-ray images of the Sun's corona, coronal holes show up as dark areas because they are regions where the corona is much more tenuous than normal. In the coronal holes, the corona does not show up in X rays and the photosphere below is very dark at X-ray wavelengths. Gas density in the coronal holes is typically about one-tenth the density of the normal portions of the corona. Near the north and south poles of the Sun, coronal holes tend to be relatively stable. Near the equatorial and midlatitude regions of the Sun, coronal holes are less stable. Coronal hole activity varies with the Sun's magnetic activity cycle. Coronal holes are therefore in some way related to the Sun's magnetic field. The largest coronal holes, which are a few hundred thousand kilometers in diameter, can last for months. More typical coronal holes are tens of thousands kilometers in diameter. These smaller coronal holes typically last only for hours rather than months.

Most of the corona contains coronal loops, which are magnetic field structures. The solar magnetic field lines come up from the lower layers of the Sun, loop into the corona, then flow back down into the solar interior. Charged particles, such as protons and electrons, in strong magnetic fields generally travel in spiral paths around the magnetic field lines. The magnetic forces do not allow these particles to travel across the magnetic field lines. Solar material is a plasma in which electrons are separated from the atomic nuclei; it is composed of charged particles and flows along these coronal loops.

Coronal loops do not exist in coronal holes. In coronal holes the magnetic field lines from the Sun's interior do not loop back into the Sun. They extend outward into interplanetary space. In the coronal holes, solar material moves upward from the Sun's lower layers along these magnetic field lines. Rather than falling back down into the Sun, this material—which is mostly protons (hydrogen nuclei) and electrons with occasional heavier, ionized atoms—streams out into interplanetary space and leaves a low-density coronal hole. Coronal holes therefore contain solar matter flowing into interplanetary space and are a major source of the solar wind.

The solar wind consists of charged particles from the Sun flowing outward into space. Were it not continually replenished from the lower layers of the Sun, the solar wind would evaporate the corona in about a day. A few billion kilograms of solar material flow outward in the solar wind every second. The Sun permanently loses this mass. It would seem that the Sun might quickly evaporate from the cumulative effect of the coronal holes and solar wind, but the Sun is very much more massive than the Earth. Hence, in the nearly five billion years of its existence, the Sun has lost less than one-tenth of 1 percent of its mass to the outflow of the solar wind.



These huge loops of solar plasma above the Sun's photosphere extend many times Earth's diameter. (NASA)

Coronal mass ejections occur when coronal loops break. Normally solar material flows along the coronal loops from the solar interior, into the corona, and back into the interior. However, occasionally the magnetic field lines in a coronal loop break. When this happens, the loop more or less explodes. The solar material in the loop is no longer confined by the magnetic field; it shoots outward into space. These events are called coronal mass ejections, or CMEs.

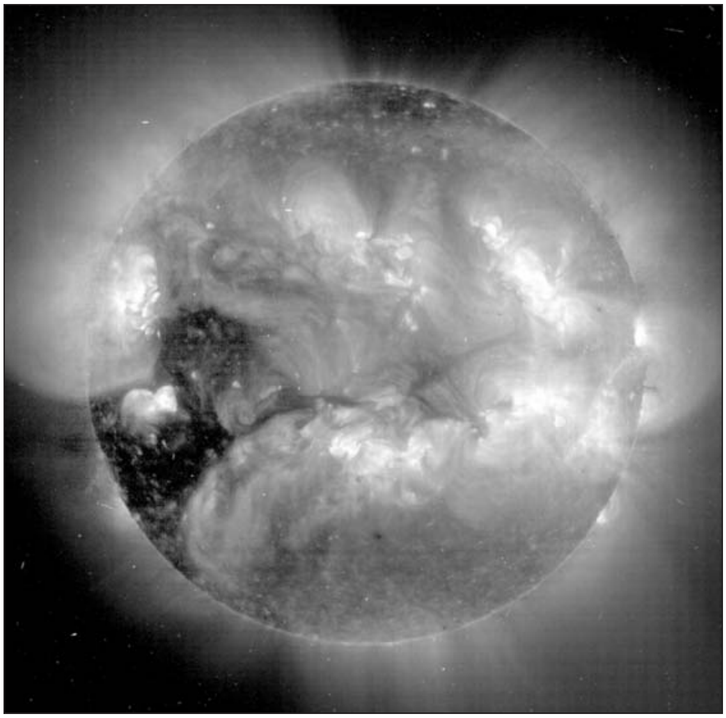
A typical CME flings about 10^{12} to 10^{13} kilograms of solar material into space. Typically a CME releases 10^{24} to 10^{25} joules of energy. Because CMEs are related to the Sun's magnetic field, they occur more frequently during the maxima of the solar activity (sunspot) cycle. CMEs happen as often as a few times a day. When a CME is pointed toward Earth, the resulting magnetic storm can severely disrupt long-distance communications on Earth and cause increased auroral activity.

KNOWLEDGE GAINED

In contrast to the Sun's photosphere, the spectrum of the Sun's corona contains emission lines, which form from a hot, thin gas. Many of the emission lines in the Sun's coronal spectrum are lines not visible from Earth. Astronomers originally thought these lines might be a new element, but they turned out to be what scientists call forbidden lines. The existence of these forbidden lines was an early clue to the extremely low density of the corona.

Although the corona can be studied optically from the ground during eclipses, much of our knowledge of the corona, coronal holes, and coronal mass ejections comes from satellites, particularly those equipped to observe the Sun at extreme ultraviolet and X-ray wavelengths as well as optically. Such studies started in earnest in the 1970's using X-ray telescopes on the Skylab mission. During this crewed mission, solar astronomers first noticed the connection between coronal holes and the solar wind.

Other satellites have been important to understanding coronal phenomena. The Japanese Yohkoh satellite was launched in 1991 and fell back to Earth in 2005. Yohkoh for the first time allowed daily images of the corona allowing solar astronomers to study rapid changes in the coronal structure. The joint European and National Aeronautics and Space Administration (NASA) Solar and Heliospheric Observatory, SOHO, mission launched in 1995, and the NASA Transition Region and Coronal Explorer (TRACE) mission, launched in 1998, also made many contributions to our understanding of coronal and other solar phenomena. SOHO discovered a magnetic carpet on the Sun's surface that plays a major role in providing the energy needed to heat the corona. TRACE takes very high-resolution, extreme ultraviolet images of the corona. Although TRACE can image only a small region of the corona at one time, it does allow very detailed studies of coronal phenomena.



The dark areas on the left of this image are coronal holes. (SOHO/NASA)

CONTEXT

Coronal mass ejections as well as variations in the solar wind can affect Earth, causing geomagnetic storms. These geomagnetic storms are often referred to as space weather. The aurora borealis and aurora australis, also known as the northern and southern lights, are caused by these geomagnetic storms. Hence they are more likely to be visible when a coronal mass ejection reaches Earth. Other geomagnetic effects are less benign. Earth's upper atmosphere expands a little and disrupts long-distance radio communications that depend on radio waves reflecting off the ionosphere or communication satellites. The expanded upper atmosphere can cause some friction on low-Earth-orbit satellites leading to orbital decay and eventually falling back to Earth. Geomagnetic storms caused by coronal mass ejections can also disrupt the electrical power grid.

Coronal holes and coronal mass ejections are part of the complex magnetic phenomena of the Sun's corona. Coronal holes seem to play a still poorly understood role in the Sun's magnetic activity cycle. These phenomena do not exist in isolation. The corona and its magnetic fields interact with the Sun's chromosphere, photosphere, interior, and their magnetic fields. Via the solar wind, originating in coronal holes, and coronal mass ejections, the Sun's corona also interacts with Earth. To understand any facet of this complex Sun-Earth system completely, astronomers need to understand all the other facets.

Paul A. Heckert

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Chapter 16 of this readable introductory astronomy textbook covers the Sun. It has a good discussion of coronal holes, coronal mass ejections, and their role in producing the solar wind.

Frazier, Kendrick. *Our Turbulent Sun*. Englewood Cliffs, N.J.: Prentice-Hall, 1980. Provides a good, readable account of our knowledge of the Sun through 1979.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. Chapter 16 of this introductory astronomy textbook is a complete, readable overview of our knowledge of the Sun, including the corona.

Golub, Leon, and Jay M. Pasachoff. *Nearest Star: The Surprising Science of Our Sun*. Cambridge, Mass.: Harvard University Press, 2001. This well-written book gives a detailed summary of our knowledge of the Sun.

Heckert, Paul A. "Solar and Heliospheric Observatory." In *USA in Space*. 3d ed. Edited by Russell Tobias and David G. Fisher. Pasadena, Calif.: Salem Press, 2006. This article describes the SOHO solar observatory mission, which was used to study the Sun's outer layers, including the corona. This mission helped us understand the Sun's magnetic activity and why the temperature increases in the chromosphere and corona. It also observed and returned data on many coronal mass ejections.

Hester, Jeff, et al. *Twenty-first Century Astronomy*. New York: W. W. Norton, 2007. Chapter 13 of this readable astronomy textbook is about the Sun.

Morrison, David, Sidney Wolf, and Andrew Fraknoi. *Abell's Exploration of the Universe*. Philadelphia: Saunders College Publishing, 1995. The Sun is covered in chapter 26 of this classic astronomy textbook.

Schrijver, Carolus J. "The Science Behind the Solar Corona." *Sky and Telescope* 111, no. 4 (April, 2006): 28-33. A good article discussing the latest understanding of the solar corona.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. Cambridge, England: Cambridge University Press, 2002. A well-written introductory astronomy textbook. Chapter 12 is an overview of the Sun.

Zeilik, Michael, and Stephen A. Gregory. *Introductory Astronomy and Astrophysics*. 4th ed. Fort Worth, Tex.: Saunders College Publishing, 1998. A textbook designed for undergraduate physics or astronomy majors, it goes into more mathematical depth than most introductory astronomy textbooks. Chapter 10 covers the Sun, including the corona.

See also: Earth-Sun Relations; Earth's Magnetic Field: Origins; Earth's Magnetic Field at Present; Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots.

Cosmic Rays

Category: The Cosmological Context

High-energy cosmic rays are samples of material from outside the solar system. Elemental and isotopic compositions of the cosmic rays constrain models for element production in a variety of astrophysical sources.

OVERVIEW

Cosmic rays are charged particles, electrons, and positively charged ions ranging from protons to the heaviest elements, which arrive at Earth from space. About 98 percent of cosmic rays are positively charged nuclei, with most of the remainder being negatively charged electrons. Although some of the lowest-energy cosmic rays are particles emitted by the Sun, most cosmic rays are too energetic to be confined to the solar system and are samples of material from other parts of the galaxy. Because cosmic rays are charged particles, their paths from their sources to Earth are bent by magnetic fields in the galaxy. As a result, traditional astronomy in which electromagnetic radiation intercepted by a detector, such as a telescope, is traced back in a straight line to its source is not possible with cosmic rays. Nevertheless, cosmic rays provide important clues to the processes that occur in stars, supernovae, and other astrophysical sources.

Determination of the composition of cosmic rays permits comparison to the composition of the Earth, lunar samples returned by the Apollo missions, meteorites, and the Sun. This allows processes by which elements are produced within stars to be examined and compared to theoretical models for nucleosynthesis.

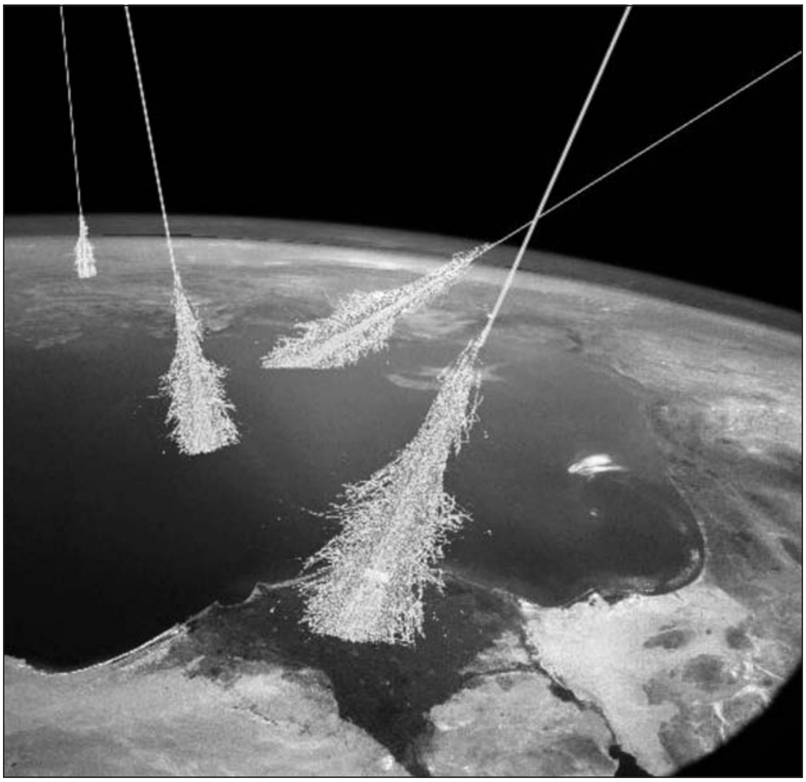
The nucleus of each element has a unique charge, so methods of determining the composition of cosmic rays require a measurement of charge on each individual cosmic-ray particle. Generally, these techniques require two independent measurements. The first measurement might determine the rate at which cosmic rays lose energy in traversing the detector. This rate of energy loss is proportional to the square of the charge-to-speed ratio of the particle. A second measurement might then determine the velocity, or some other property that depends on velocity, in a manner different from that for determining the rate of energy loss. From these two measurements, the charge can be determined.

A number of innovative charge-measurement detectors have been developed. These detectors can be divided into three general categories: recording detectors, such as photographic emulsions; visual detectors, such as cloud chambers; and electronic detectors, such as Geiger-Müller counters.

In the late 1940's, groups of cosmic-ray investigators from the University of Minnesota and the University of Rochester employed photographic emulsions carried to high altitudes by balloons to determine the charge and energy of the cosmic rays. High altitudes, frequently above 27,000 meters, were required because collisions between incoming cosmic rays and air molecules can cause cosmic rays to fragment into several lighter nuclei, thus altering their composition. At high altitudes, the probability of such a collision is low. Therefore, balloon detectors measure the primary composition of the particles in space. These early experiments demonstrated that, of the nuclei in cosmic rays, about 87 percent are hydrogen, or protons; 12 percent are helium; and the remaining 1 percent are heavier nuclei. It is the composition of these heavier nuclei that contains the clues to the nucleosynthetic processes.

Following the initial discovery of heavy nuclei among cosmic rays, the emphasis in cosmic-ray research shifted to the determination of the charge spectrum, or relative abundances, of each element. Early experiments made use of the Earth's magnetic field as a velocity selector. Paths of charged particles are bent when they encounter a magnetic field, so only particles exceeding a given cutoff energy can penetrate through a region of given magnetic field intensity. The magnetic field of the Earth is so strong near the equator that only cosmic-ray particles with velocities very close to the speed of light can penetrate. Thus, for those cosmic rays detected near the equator, the magnetic cutoff identifies the velocity to be approximately (just under) the speed of light. A single measurement of the rate of energy loss for these particles provides a measurement of their charge.

These early experiments indicated the difficulty of de-



Showers of cosmic rays fall to Earth after striking the atmosphere in this artist's conception. (Simon Swordy/NASA) tection of the heavy nuclei among cosmic rays. A one-square-meter detector placed in space, above the Earth's atmosphere and outside the Earth's magnetic field, would register several hundred non-solar protons per second and about one-seventh that number of helium nuclei. However, only one or two nuclei heavier than carbon would be measured every second, and the detector would register a single iron nucleus every fifteen seconds. To observe a single lead nucleus would require several months of detector operation. Cosmic-ray astrophysicists recognized that large detectors with long exposure times would be required to determine accurately the composition of the heavy cosmic rays.

In 1956, Frank McDonald, a physicist at Iowa State University, developed a combination of two electronic detectors—a scintillation counter and a Cherenkov counter—to determine the charge and velocity of the cosmic rays. This combination of detectors provided good measurements of the elemental abundances for elements up to iron. Elements heavier than iron were so rare that their identification required a new technique. In the mid-1960's, Robert Fleischer, Buford Price, and Robert Walker, researchers at the General Electric Research and Development Center, found that trails of ionizing particles were recorded within certain types of plastics and that these trails could be revealed later by etching the plastic in an appropriate chemical agent. They demonstrated that if the rate at which the trail was etching as well as the total etchable length were both measured, the charge and energy of the particle could be determined. Balloon flights with these plastic detectors provided information on the composition of the heavier elements in cosmic rays.

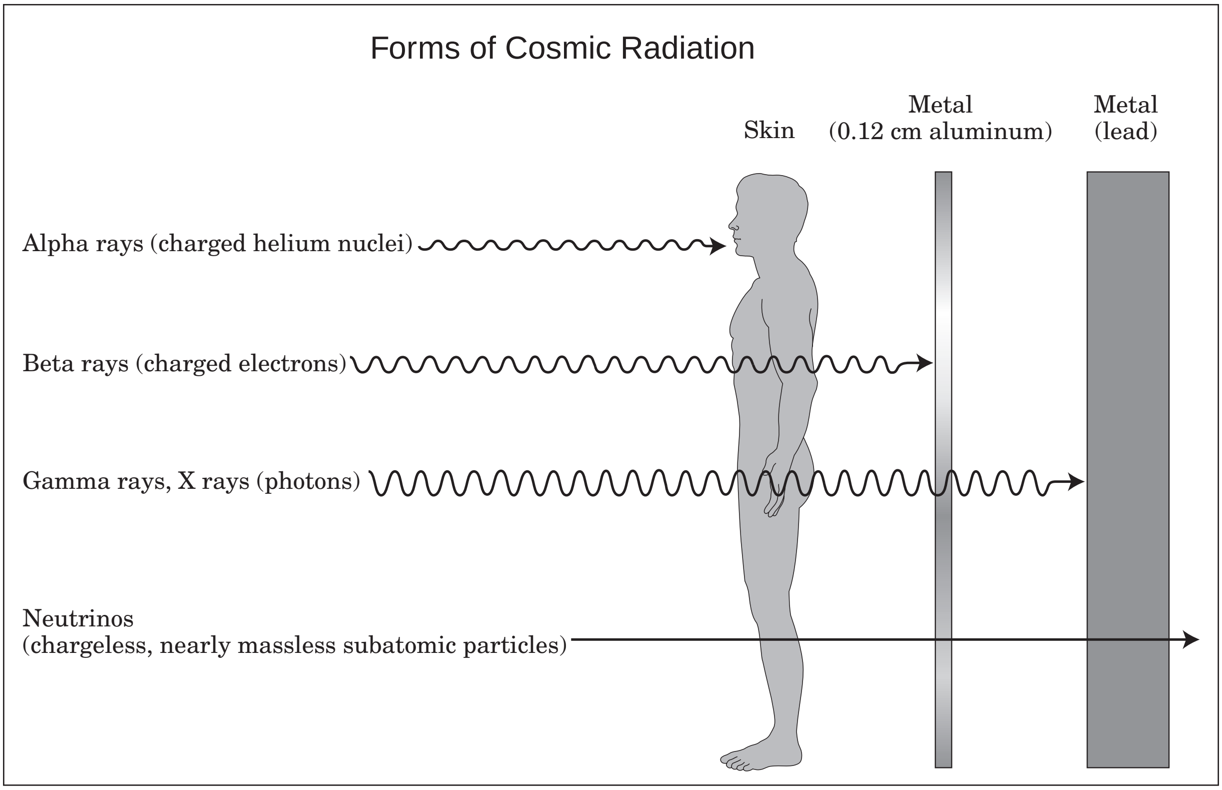
In the 1970's, cosmic-ray researchers employed orbiting Earth satellites to increase the duration of their measurements. Large plastic detectors were flown for several months on the National Aeronautics and Space Administration's (NASA's) Skylab space station. The IMP-7 and IMP-8 satellites, launched in 1972 and 1973, respectively, provided good measurements on the isotopic composition of lighter cosmic rays. In 1978, the third High-Energy Astronomical Observatory (HEAO 3), carrying a 6-square-meter electronic detector, provided high-quality measurements of the abundances of nuclei up to bismuth. Cosmic-ray studies continue at polar stations, with small university-based instruments examining the space environment surrounding Earth and with national and international spacecraft journeying throughout the solar system.

APPLICATIONS

Astrophysicists generally believe that the only elements present in the early universe were hydrogen, helium, and perhaps small amounts of lithium, beryllium, and boron. Most of the elements now present were produced in stellar explosions, or by nucleosynthesis in stars. Theoretical calculations show that the elemental and isotopic abundances produced depend on the particular conditions of the nucleosynthetic event. Thus, abundances of the elements and isotopes of material from outside Earth's solar system might be different from those of solar-system material, and those differences would provide clues to the differing nucleosynthetic conditions at those sites. This comparison requires a knowledge of the cosmic-ray composition at the source.

The composition of the cosmic rays can be altered during their journey through space to Earth. Radioactive decay will remove those radioactive elements with short half-lives, compared to the time it took for cosmic rays to reach Earth. Collisions between cosmic rays and interstellar gas will cause fragmentation of some of cosmic rays.

Measurements of cosmic-ray composition provide clues to the "age" of cosmic rays, that is, the duration of their journey through space. The light nuclei—lithium, beryllium, and boron—are much more abundant in cosmic rays than in solar-system materials. These excess light nuclei are believed to have been produced by spallation, or collisional fragmentation with interstellar gas atoms. Since the abundance of interstellar gas atoms is known from other astronomical measurements, the amount of excess light elements provides a measure of the duration of the cosmic-ray journey. Astrophysicists indicate that cosmic rays currently arriving at Earth began their journey about 10 million



Different forms of cosmic radiation can penetrate different forms of matter: Alpha rays cannot penetrate skin; beta rays can penetrate skin but not metal; gamma rays can penetrate both but are stopped by lead; and neutrinos, chargeless, nearly massless particles, can penetrate even lead, making them extremely difficult to detect. Although they interact very little with matter, neutrinos are believed to be produced in the nuclear reactions at the core of the Sun and other stars and may constitute a large portion of the "missing mass" of the universe.

years ago. Since the solar system formed about 4.5 billion years ago, the cosmic rays may be sampling a much younger type of material than the solar system.

Once the age of cosmic rays is known, the abundances of heavier elements detected at Earth can be corrected back to the source by removing the spallation contribution. Generally, elemental composition of cosmic rays is similar to that of the solar system, but the differences provide clues to differences in the nucleosynthetic processes.

The largest difference in the heavy element composition is for the isotope neon 22, which is four times more abundant relative to the other neon isotopes in cosmic rays than in solar-system matter. Isotopic measurements also show excesses of magnesium 25, magnesium 26, silicon 29, and silicon 30 in cosmic rays when compared to solar-system matter. These latter discrepancies could be explained if the cosmic-ray sources were stars with initial abundances of carbon, nitrogen, and oxygen about twice that seen in Earth's Sun. Nevertheless, even this alteration of the stellar composition cannot explain the unusually high abundance of neon 22.

Abundances of the heavier elements may provide clues to the site of the production of cosmic rays. Astrophysicists have identified several different nucleosynthetic processes. Two major ones both proceed by the addition of neutrons to light target nuclei. In the *s*-process, the time between successive neutron capture events is long enough that the new nucleus can be transformed (beta decay) to a stable nucleus before the next capture. This process occurs in the interiors of stars. In the *r*-process, neutron capture proceeds so rapidly that beta decay is not possible between individual capture events, leading to production of more neutron-rich elements. This process is believed to occur in explosive processes such as supernovae. Only the *r*-process can produce elements heavier than bismuth.

Since astrophysicists have suggested that supernovae are a likely source of cosmic rays, they would be expected to contain the *r*-process elemental and isotopic abundance signatures. The presence of elements heavier than bismuth in cosmic rays would suggest an *r*-process origin. Thus far, experimental results are ambiguous. Rare events, possibly attributable to elements heavier than bismuth, were reported from balloon flights carrying photographic emulsions and plastic detectors. However, the HEAO 3 detected no such events. Because of the scarcity of these heavy elements, longer-duration, large-area cosmic-ray detectors will be required to resolve the question.

Elemental and isotopic measurements on cosmic rays indicate that their sources differ in significant ways from the source of solar-system material. Because of these differences, more precise measurements of elemental and isotopic compositions of the heavier elements are required for detailed comparisons to the nucleosynthetic models.

CONTEXT

Formulation of a detailed model of the nucleosynthesis of the heavy elements by Geoffrey Burbidge, Margaret Burbidge, William A. Fowler, and Fred Hoyle in 1957 provided predictions of the elemental and isotopic abundances expected from the *r*-process and *s*-process. This information, coupled with the discovery of heavy elements in cosmic rays in the late 1940's, suggested comparison of cosmic-ray composition with the predictions of the nucleosynthetic models. Rapid advances in electronic detectors in the 1950's made such comparisons possible, but the limited flight duration of high-altitude balloons restricted the number of elements that could be measured because of the low abundance of heavy elements. The use of Earth satellites in the 1970's significantly increased the duration of cosmic-ray composition experiments. Nevertheless, even these long-duration satellite experiments were inadequate to answer the question of the abundance of heavy elements in the cosmic rays.

Development of high-resolution electronic detectors, permitting high-quality determinations of isotopic composition, showed significant differences between the neon, magnesium, and silicon isotopic abundances in cosmic rays and solar-system matter. Advances in modeling of the nuclear processes in stellar interiors allowed astrophysicists to calculate that most of these discrepancies were consistent with stellar nucleosynthesis, with carbon, nitrogen, and oxygen abundances approximately double that of Earth. Long-duration, large-area cosmic-ray detectors, possibly on a space station, will be required to determine the abundances of elements heavier than bismuth, allowing direct comparison of the cosmic-ray composition with that expected for *r*-process nucleosynthesis in supernovae, which are suggested as the cosmic-ray source.

George J. Flynn

FURTHER READING

Cronin, James W., Thomas K. Gaisser, and Simon P. Swordy. “Cosmic Rays at the Energy Frontier.” *Scientific American* 276, no. 1 (January, 1997): 44-50. A technical but thorough examination of our understanding of high-energy cosmic rays.

Foerstner, Abigail. *James Van Allen: The First Eight Billion Miles*. Iowa City: University of Iowa Press, 2007. An engaging portrait of the legendary physicist James Van Allen. Discusses his contributions to the World War II effort as well as the advancement of studies of Earth's geomagnetic environment, his early efforts to study cosmic rays using balloon-launched rockets, the Explorer 1 story, and Van Allen's continuing participation in studying space physics until his death in 2006.

Friedlander, Michael W. *Cosmic Rays*. Cambridge, Mass.: Harvard University Press, 1989. Well-illustrated account of the history of cosmic-ray astronomy. Deals with methods of detection, elemental and isotopic composition, and implications for the cosmic-ray sources.

Ginzburg, V. L., and S. I. Syrovatskii. *The Origin of Cosmic Rays*. New York: Macmillan, 1964. A technical account of cosmic-ray astrophysics. Includes a good discussion of how the light element abundances provide an age for the cosmic rays. Suitable for college physics students.

Pomerantz, Martin A. *Cosmic Rays*. New York: Van Nostrand Reinhold, 1971. A classic text suitable for readers with only an introductory physics background. Describes the interactions of cosmic rays with matter and how these interactions are used to detect and determine the properties of the cosmic rays. Well illustrated.

Rossi, Bruno. *Cosmic Rays*. New York: McGraw-Hill, 1964. A firsthand account by one of the pioneers of cosmic-ray physics. Describes how cosmic rays are detected and discusses ideas about their origins.

Van Allen, James A. *The Magnetospheres of Eight Planets and the Moon*. Oslo, Norway: Norwegian Academy of Science and Letters, 1990. A technical summary of all major magnetic structures in the solar system, written by the famous and prolific researcher after whom the Van Allen belts are named.

Wefel, John P. "Matter from Outside Our Solar System: New Insights. Part 1, The Astrophysical Framework." *The Physics Teacher* 20 (April, 1982): 222-229. Presentation of cosmic rays at a level useful for both high school and undergraduate physics.

_______. "Matter from Outside Our Solar System: New Insights. Part 2, Experimental Measurements and Interpretation." *The Physics Teacher* 20 (May, 1982): 289-297. Discusses the history of cosmic-ray physics, the mechanisms of nucleosynthesis, the construction of cosmic-ray detectors, and the implications of the composition on the sources. Well illustrated. Suitable for high school science students.

See also: Big Bang; Cosmology; Earth's Magnetosphere; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; Extraterrestrial Life in the Solar System; General Relativity; Interplanetary Environment; Interstellar Clouds and the Interstellar Medium; Lunar Rocks; Milky Way; Radio Astronomy; Solar Flares; Solar Magnetic Field; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Universe: Evolution; Universe: Expansion; Universe: Structure; X-Ray and Gamma-Ray Astronomy.

Cosmology

Category: The Cosmological Context

Cosmology is the study of the structure and evolution of the universe, including the eventual development of our solar system. It brings together many branches of astronomy and physics, including general relativity and high-energy particle physics.

OVERVIEW

Originally, cosmology was a branch of philosophy devoted to understanding the nature of reality and the origin and structure of everything that exists. With the growth of astrophysics during the nineteenth and twentieth centuries, cosmology rapidly became a major area of research in astronomy and physics, and its focus narrowed to the origin and evolution of energy and matter in the universe as a whole. Cosmology today is concerned with the large-scale structure of the universe, including the distribution of billions of galaxies and galaxy clusters throughout space and time. Cosmologists use physical laws to derive mathematical models of the early universe to within 10^{-43} second of its beginning. They also extrapolate physical processes into the distant future to predict the future of the universe and its contents. Nevertheless, modern cosmology still retains some philosophical qualities.

Modern cosmology has its basis in Albert Einstein's general theory of relativity, which he published in 1915. In 1922, the Russian mathematical physicist Alexander Alexandrovich Friedmann derived two types of solutions to the field equations of general relativity in which the universe initially expands with time. In one type (called "open"), the universe continues to expand forever. In the other type (called "closed"), the universe expands to some maximum size, after which it contracts. In 1927, the Belgian priest and cosmologist Abbé Georges Lemaître independently derived the same two solutions to the general relativistic field equations and went on to speculate about the cause of the expansion; he suggested the universe began from a compact, dense initial state—the "primeval atom"—which disintegrated and dispersed into all the atoms in the universe today.

Edwin P. Hubble presented observational evidence that the universe actually is (or at least appears to be) expanding, when in 1929 he published his research that showed a linear correlation between a galaxy's distance from Earth and the redshift of its spectrum. The cosmological explanation for the redshift is that, as the universe expands, the wavelengths of all electromagnetic radiation are stretched. The color red has the longest wavelength of all components of visible light. When wavelengths of visible light are stretched and shifted toward the red, they are said to be redshifted. The use of the term "redshift" has also been extended to refer to a shift toward longer wavelengths in any region of the electromagnetic spectrum.

The generally accepted cosmological theory describing the origin and expansion of the universe is the big bang theory, which can be traced back to the proposal by Lemaître of a compact, dense initial state that somehow exploded. The physicist George Gamow expanded on Lemaître's idea in the 1930's and 1940's, making it physically more rigorous and incorporating new work in nuclear physics. The name "big bang" was coined by Fred Hoyle, one of the developers of the rival steady state, continuous creation theory as a derogatory term for an explosive ori-

An image from the Hubble Space Telescope's Ultra Deep Field shows galaxies going back billions of light-years, to within a billion years of the big bang. (NASA/ESA/S. Beckwith, STScl) gin, but it rapidly caught on and was adopted by proponents and opponents alike. By the mid- to late 1960's, observational evidence had eliminated competing theories, and the big bang became generally accepted. Throughout the remainder of the twentieth century and into the twenty-first century, the big bang has been further modified and refined, and now it incorporates the latest work in high-energy particle physics.

The big bang theory maintains that space, time, energy, and matter were all created from a primordial explosion, now dated at about 13 to 14 billion years ago. Space expanded isotropically, and as it expanded, the universe cooled. Modern physical theories can be used to trace the development of the universe back to 10^{-43} second after the big bang, when the temperature was 10^{32} kelvins, but before that time the theories break down. At the even higher temperatures prior to that time, scientists think all four fundamental forces of nature—gravity, strong nuclear, weak nuclear, and electromagnetism—were all unified as one force, indistinguishable from one another. As the universe expanded and cooled, the forces gradually separated from each other one by one—first gravity, then strong nuclear, and finally weak nuclear from electromagnetism at about 10^{-10} second and 10^{15} kelvins.

The early universe was a hot, dense "soup" of interacting high-energy photons and subatomic particles, with energy and mass being transformed back and forth. When two photons with enough energy collided, their energy could be converted to mass, producing a matter-antimatter pair of particles in a process called pair production. When a particle and its antiparticle collided, they mutually annihilated each other with their mass being converted into two high-energy gamma-ray photons.

As the universe expanded and cooled, eventually photons did not have enough energy to produce any more particle-antiparticle pairs. Thereafter, the particles and their antiparticles collided and mutually annihilated. Since equal numbers of particles and antiparticles had been created, they all should have annihilated each other and the universe today would be devoid of matter and antimatter. Alternatively, some segregation process might have separated the matter and antimatter into distinct regions, and the universe today would consist of equal but separate concentrations of matter and antimatter. However, the universe today appears to be composed almost entirely of matter throughout. An asymmetry in the weak nuclear force seems to provide a way for antimatter but not matter to decay, so there was a slight excess of matter particles by about one in a billion. By the end of a few seconds, the temperature had dropped to several billion kelvins, and particle creation and annihilation had ceased. The small excess of matter that survived is the matter of the universe today.

Among the particles to survive were quarks. They combined to form protons and neutrons; a proton is two "up" quarks and one "down" quark, while a neutron is one "up" quark and two "down" quarks. Protons and neutrons are the particles that make up atomic nuclei. Single protons are hydrogen nuclei, and they combined with neutrons to form nuclei of other light elements such as deuterium (also called heavy hydrogen, with one proton and one neutron), helium (two protons and one or two neutrons), and small amounts of lithium (three protons and three or four neutrons) and beryllium (four protons and three neutrons). However, after about fifteen minutes, the temperature had dropped to a few hundred million kelvins, too cool for further nucleosynthesis. Heavier nuclei would be formed much later, in nuclear fusion reactions in stars.

The early universe was dominated by electromagnetic radiation; that is, the spatial density of electromagnetic energy was greater than the spatial density of matter. Both densities decreased as the universe expanded, but the electromagnetic energy density decreased faster than the matter density. Several thousand years after the big bang, the electromagnetic energy density dropped below the matter density at a time referred to as the crossover time. After that, the universe was dominated by matter, since its density was greater.

Around 300,000 to 500,000 years after the big bang, the temperature had dropped to about 3,000 kelvins. Electrons could then join with protons (bare hydrogen nuclei) to form electrically neutral hydrogen atoms. Free electrons are very effective at scattering photons, but electrons in atoms are not able to do so. As a result, the universe changed from being very opaque to becoming transparent to electromagnetic radiation. Photons could now travel freely through the universe. This was the source of the cosmic microwave background radiation currently observed at a temperature of about 3 kelvins, because the wavelength of the electromagnetic radiation has been greatly stretched (redshifted) by the expansion of the universe.

Within the first few hundred million years, matter clumped together by gravitational attraction to form protogalaxies or pregalactic fragments, and within them further gravitational clumping formed the first stars. These protogalaxies were relatively small, but through mergers they developed into larger systems, the galaxies of the universe today. Galaxies range in size from dwarfs, containing a few tens of millions of stars, to giants, with more than ten trillion stars. Galaxies are not distributed randomly through space but are grouped into galaxy clusters; poor clusters contain only a few tens of galaxies, while rich clusters have thousands of members. Galaxy clusters in turn are grouped into superclusters. Between them are large, nearly empty regions called voids.

Within galaxies, stars form from clouds of gas and dust called nebulae. Stars initially heat up and begin to shine by gravitational contraction, but this is a relatively brief stage in the life cycle of a star. During most of their energy-producing lives, stars generate energy by nuclear fusion reactions in which lighter atomic nuclei are fused into heavier nuclei with the release of energy. Stars with many times the Sun's mass can synthesize nuclei as heavy as iron in their interiors. When they explode as Type II supernovae at the end of their energy-producing lives, the tremendous energy released synthesizes nuclei heavier than iron. The explosion disperses the elements the star formed during its life out into space, there to enrich the nebulae from which new generations of stars form.

It had been assumed that in the future, the expansion of the universe will slow down, due to the gravitational attraction between galaxies. The only question has been how rapidly the expansion is decelerating. If the deceleration were small, the universe would expand forever, at a gradually decreasing rate; if the deceleration were large enough, someday the universe would stop expanding and begin to contract at ever-increasing speed. Beginning in the 1990's, astronomers tried to determine how much the universe is slowing down by measuring the expansion rate at great distances (and hence at great times in the past) and comparing it to the expansion rate at smaller distances (and more recent times). Contrary to expectations, they found the expansion rate in the past was slower than it is now, indicating that the expansion of the universe is accelerating. The cause for this acceleration is unknown, but it is called "dark energy." If the acceleration continues, the distances between galaxy clusters will grow at an ever-increasing rate. Eventually, all the matter in galaxies will be processed into stars, all the stars will use up their sources of energy and go out, and the universe will grow cold and dark.

METHODS OF STUDY

Cosmology is studied both theoretically and observationally, the two complementing each other. New observations need to be interpreted by theories, and theories need to be confirmed by further observations. Cosmological observations are now made over the entire electromagnetic spectrum, from high-energy, short-wavelength gamma rays, through X rays, ultraviolet, visible light, infrared, microwaves, to long-wavelength radio waves. Not all electromagnetic radiation penetrates Earth's atmosphere, so parts of the electromagnetic spectrum must be observed from satellites above Earth orbit, above the atmosphere.

The speed at which all forms of electromagnetic radiation travel through a vacuum is exceedingly fast but finite, being very close to 300,000 kilometers per second. Therefore, looking out to greater distances means looking further back in time. When astronomers look at Earth's moon, at a distance of about 400,000 kilometers, they see it as it was about 1.3 seconds earlier. When one observes the Sun, at a distance of about 150 million kilometers, it appears as it was 8 minutes and 20 seconds earlier. The nearest star system outside our solar system, the Alpha Centauri system, is at a distance of 4.3 light-years, meaning that we see that system as it was 4.3 years ago. Distant galaxies are billions of light-years away, so we see them as they were billions of years ago. In this manner it is possible to observe the early universe and its contents by observing at larger and larger distances.

There are two primary observational anchors in modern cosmology. The first is the Hubble law, the relationship between galaxy redshifts and distances, first discovered by Edwin Hubble. This provides the basic observational evidence that the universe is expanding. Other explanations for the redshifts of galaxy spectra have been proposed. For example, the "tired light" hypothesis posits that photons lose energy and hence are shifted to lower frequencies and longer wavelengths as they travel immense distances. However, none of these alternative explanations fits the observed data, with a minimum of extra assumptions, as well as the expanding universe does—which explains the redshifts of distant objects as due to the stretching of wavelengths of electromagnetic radiation as the universe expands.

The second observational anchor in cosmology is the cosmic microwave background (CMB) radiation, which is the firmest observational evidence supporting a big bang origin to the expanding universe. It was first detected accidentally by Arno A. Penzias and Robert W. Wilson in 1965 as part of their work on a communication satellite project at AT&T's Bell Laboratories in Holmdel, New Jersey. Using a large radio horn antenna, they found a uniform microwave background signal coming from all directions. It was identified by Robert Dicke and his colleagues at Princeton University as greatly redshifted radiation from a few hundred thousand years after the big bang, the time when the universe became transparent. Subsequent observations of it by Earth-orbiting spacecraft—the Cosmic Background Explorer (COBE) and the Wilkinson Microwave Anisotropy Probe (WMAP)—and high-altitude balloons launched in Antarctica, including Balloon Observations of Millimetric Extragalactic Radiation and Geomagnetics (BOOMERANG), showed that the radiation exactly fits a blackbody spectral curve for a temperature of 2.73 kelvins. However, it is not precisely uniform. There are temperature variations of up to about 0.00001 kelvin over areas of the sky with an angular size of about 1° of arc (about twice the apparent angular size of the Sun and full Moon as seen in our sky). These temperature variations are thought to represent slight differences in density in the early universe that ultimately produced the "lumpy" universe of clusters of galaxies that observed today.

On the theoretical side, modern cosmology draws primarily upon general relativity and high-energy particle physics. Solutions to the simplest form of the field equations of general relativity predicted an expanding universe before it was confirmed observationally. High-energy particle physics provides insights into the processes that likely occurred in the high-temperature, high-energy environment of the early universe.

Grand unified theories (GUTs) seek to unify the strong nuclear, weak nuclear, and electromagnetic forces as manifestations of a single more fundamental force, and theories of everything (TOEs) try to include gravity with the other three forces. It is thought that this unification, meaning the forces are indistinguishable from each other, occurs at extremely high temperatures and energies, the conditions that existed in the very early universe. Thus the very early universe serves as a laboratory to test such theories.

CONTEXT

The value of cosmology lies in understanding the structure and organization of the universe, where it came from, and how it will develop in the future. Cosmology gives us a perspective on our place in the universe. At the same time, cosmology provides a way to test the laws of physics on a grand universal scale.

Two of the major unsolved questions in cosmology involve the nature of dark matter and dark energy. There are numerous situations in astronomy in which a study of the dynamics of a system implies gravitational forces that far exceed what the observed mass can account for. The deficit in observed mass was originally called “missing mass.” However, the dynamical mass calculations seem reliable, so astronomers now generally use the term “dark matter,” since the mass really is not missing, it is just not observable in any part of the electromagnetic spectrum. Dark matter probably includes nonluminous ordinary matter that has not been observed yet, such as small conglomerates of nonradiating matter, black dwarfs, and black holes. However, indications are that most dark matter is much more exotic—completely unknown forms of matter that do not interact with ordinary matter except gravitationally. Possible candidates include a class of particles called WIMPs (weakly interacting massive particles) and cosmic strings (long, thin, massive lines of unbroken symmetry left over from the early universe in which the strong, weak, and electromagnetic forces remain unified).

Dark energy, which drives the acceleration of the expansion of the universe, is even more enigmatic. Mathematically, it may take the form of Einstein’s cosmological constant in the equations derived from general relativity that describe the expanding universe. Physically, its nature is completely unknown.

Various observations indicate that the geometry of the universe is almost precisely flat. That means the average density of matter and energy throughout the universe must almost exactly equal a value called the critical density, about 10^{-26} kilograms per cubic meter. Observed luminous matter accounts for about 1 percent of this. Allowing for probable nonluminous but ordinary dark matter gives about 3 percent more. It is estimated there is about 26 times more dark matter (both ordinary and exotic) than luminous matter, so exotic dark matter accounts for 23 percent of the critical density. The total for all forms of matter comes to about 27 percent; thus dark energy contributes about 73 percent of the average density of the universe. That means approximately 96 percent of the universe consists of dark matter and dark energy, about which we know virtually nothing. Only about 4 percent consists of ordinary matter, both luminous and nonluminous.

Another cosmological puzzle is that the universe seems “fine-tuned” for life. If the physical laws and constants of the universe were much different from what they are, life as we know it would be impossible. Stars and planets would not form, or they would not last long enough for life, especially intelligent life, to develop. One explanation for this fine-tuning is called the anthropic principle: the idea that the universe has to be the way it is, because otherwise we would not exist to ask about such things. However, some scientists find the odds overwhelmingly against the universe being the way it is solely by chance, proposing instead that the universe in some way may have been deliberately designed for life. To avoid the theological implications of deliberate design, other scientists suggest that the universe is just one of many alternate parallel universes, each with its own unique set of laws and constants; human beings occupy the one that allows the existence of life.

Workable GUTs and TOEs are needed to describe the very early universe. One theory for unifying all four forces requires eleven dimensions—the familiar three dimensions of space and one of time, plus seven more dimensions. The extra dimensions would be rolled up into structures too small to detect. In some versions of this theory, particles such as quarks and electrons are really multidimensional membranes wrapped around the extra dimensions. Multidimensional membranes are also called M-branes or just branes. It has even been suggested that collisions between branes lead to big bangs, with each collision and big bang creating a new universe.

Through the 1980’s, most of the parameters that characterize the universe and its expansion were very poorly known, often with more than a factor of two uncertainty. Since then, new observations have dramatically narrowed the range of uncertainty. For example, current determinations of the Hubble constant—the slope of the redshift-distance relation, which is the rate at which the universe is expanding—are generally between 65 and 75 kilometers per second per megaparsec. (A megaparsec is a million parsecs or 3,260,000 light-years.) These values mean that the average speed with which other galaxies recede from us increases by 65 to 75 kilometers per second for every million parsecs (or 3,260,000 light-years) of distance from us. This small range of values for the Hubble constant, together with the matter and energy density percentages (27 percent matter, 73 percent energy), yields a time back to the big bang of 13 to 14 billion years ago.

Cosmology started as a branch of philosophy, but it has become an integral part of astronomy and physics. It has shifted from being primarily a speculative, qualitative endeavor to becoming a precise, quantitative science. Nevertheless, many aspects of cosmology remain philosophical in scope. Cosmologists speculate whether or not our universe is the only universe, whether or not there are additional dimensions to the known universe that have yet to be discovered, and how the four principal forces of nature might have been unified in the very early universe right after the big bang. Since humans are part of the universe, it can be said that "we are the universe contemplating itself."

David Wason Hollar, Jr., and Richard R. Erickson

FURTHER READING

Bartusiak, Marcia. *Thursday's Universe*. New York: Times Books, 1986. A thorough survey of major twentieth century breakthroughs and theories in astrophysics. Discusses the development of major cosmological principles and the people behind these ideas. Big bang cosmology is described clearly from very early stages of the universe to the distant future universe.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A well-written college-level textbook for introductory astronomy courses. Two chapters provide a thorough discussion of cosmology.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Several chapters contain material concerning cosmology.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook. Two chapters deal with cosmology.

Guth, Alan H., and Paul J. Steinhardt. "The Inflationary Universe." *Scientific American* 250 (May, 1984): 116-129. This general review article is an excellent presentation of big bang cosmology with the inclusion of inflation, in which the very early universe briefly undergoes extremely rapid expansion.

Hawking, Stephen W. *A Brief History of Time: From the Big Bang to Black Holes*. New York: Bantam Books, 1988. This enormously popular best seller presents a clear, outstanding discussion of cosmology. Hawking describes the evolution of the universe and grand unified theories, among other topics.

Kippenhahn, Rudolf. *Light from the Depths of Time*. New York: Springer, 1987. An exciting description of cosmology and the universe. Kippenhahn, an astrophysicist, covers the history of cosmological thought in the twentieth century and outlines the evolution of the universe and basic cosmological principles using humorous fictional characters.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses, divided into many short units on specific topics. Several units provide a thorough discussion of cosmology.

Seielstad, George A. *At the Heart of the Web: The Inevitable Genesis of Intelligent Life*. Boston, Mass.: Harcourt Brace Jovanovich, 1989. An excellent survey of cosmological thought. Seielstad describes the order and evolution of the universe while stressing the anthropic principle.

Silk, Joseph. *The Big Bang*. Rev. ed. New York: W. H. Freeman, 1989. A comprehensive, readable discussion of cosmological views on the origin and evolution of the universe.

See also: Big Bang; Cosmic Rays; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; General Relativity; Interstellar Clouds and the Interstellar Medium; Milky Way; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Universe: Evolution; Universe: Expansion; Universe: Structure.

D

Dwarf Planets

Categories: Planets and Planetology; Small Bodies

The discovery of many new bodies orbiting the Sun beyond the orbit of formerly outermost Neptune—including at least one larger than Pluto—created a crisis in astronomy. It became evident that a new definition was required to distinguish these objects from traditional planets. The term "dwarf planet" was introduced to include planetary objects smaller than planets but larger than asteroids, resulting in the demotion of Pluto from its status as a planet.

OVERVIEW

The concept of a planet has a long history, leading to a total of nine planets in the solar system until discoveries in the early twenty-first century led to new definitions that excluded Pluto. The word "planet" originates from a Greek word meaning "wanderer" and for centuries was applied to celestial objects that shifted positions relative to the "fixed" stars. In classical antiquity, seven such objects were identified and were associated with mythical gods: the Sun, the Moon, Mercury, Venus, Mars, Jupiter, and Saturn. The Latin names for the seven days of the week were based on these seven celestial deities. In Greek thought, the planets were believed to orbit the Earth along complex paths determined by a combination of circles.

During the scientific revolution of the sixteenth and seventeenth centuries, it was shown that five of the classical planets revolve around the Sun in elliptical orbits, along with the Earth-Moon system. Late in the eighteenth century, British astronomer William Herschel, aided by his sister Caroline, discovered Uranus, the first planet to be discovered with the aid of a telescope. Early in the nineteenth century, Sicilian astronomer Giuseppe Piazzi discovered what he thought was a new planet, smaller than Mercury and orbiting the Sun between Mars and Jupiter. He called it Ceres. However, when many smaller bodies with similar orbits were discovered in the next few decades, they were called asteroids, and Ceres was demoted from its status as a planet (though later reinstated as a "dwarf planet"). The asteroids are believed to be remnants from the formation of the solar system.

By the middle of the nineteenth century, investigations into slight deviations in the elliptical orbit of Uranus led to the discovery of Neptune by the German astronomer Johann Galle. Using Sir Isaac Newton's law of universal gravitation, astronomers were able to determine the masses of all but two of the eight known planets from the motions of their satellites, with Jupiter as the most massive, at 318 times the Earth's mass. Perceived deviations in the orbit of Neptune led to the discovery of Pluto in 1930 by American astronomer Clyde Tombaugh. Pluto's orbit differed from those of the other planets, with its large inclination from the ecliptic plane and its highly elliptical shape that brings it closer to the Sun than Neptune during some 20 years of its 248-year period. It was also found to be much smaller than the outer gas giant planets and to consist mostly of icy materials.

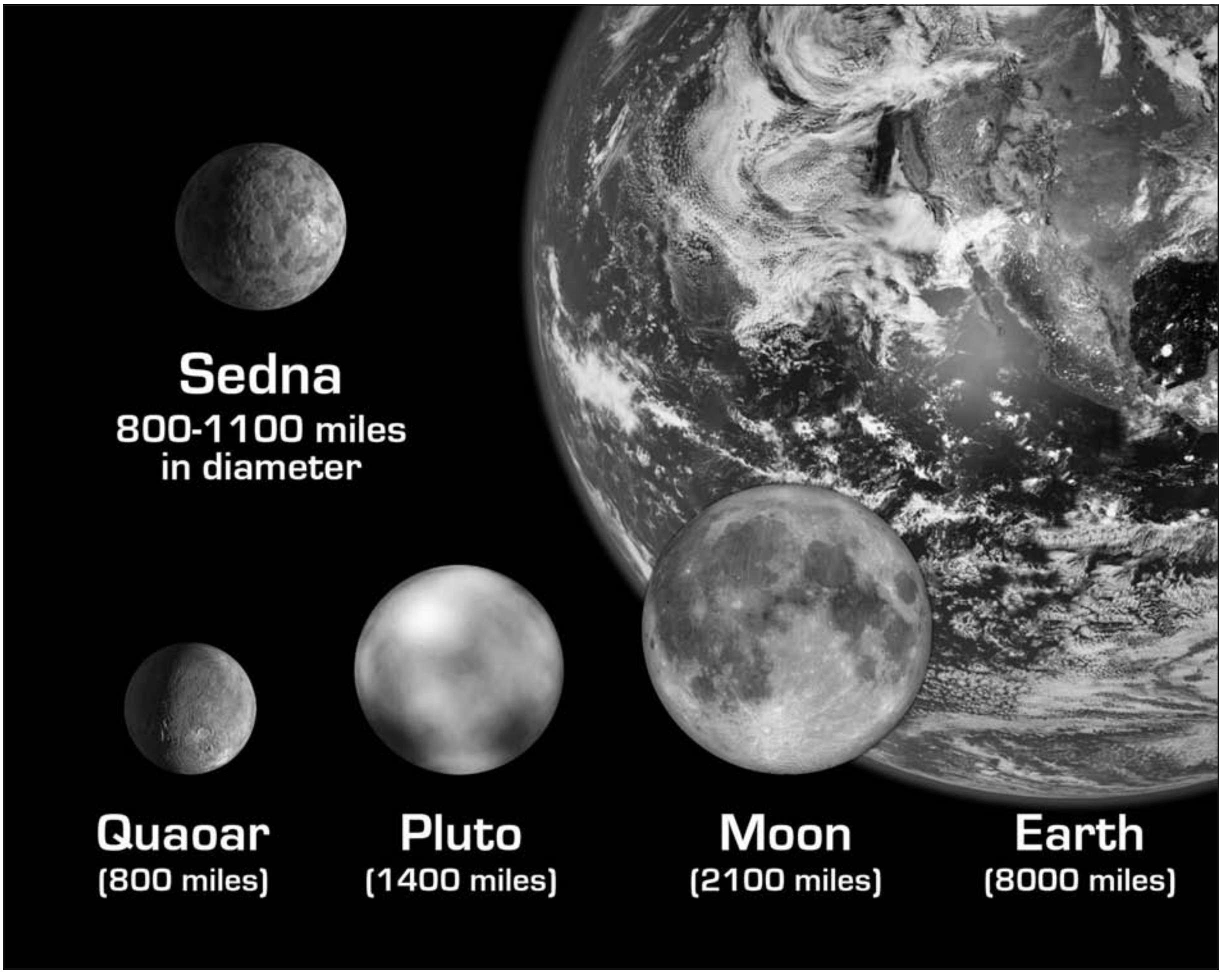
In 1977 Charles Kowal discovered a small, icy planetoid orbiting the Sun between Jupiter and Uranus, later named Chiron. In the 1990's several similar, cometlike objects were found between Jupiter and Neptune and are now called centaurs. Pluto's status as the ninth planet began to be suspect in 1978, when its satellite Charon was discovered and Pluto's mass was found to be only 0.2 percent of Earth's mass. That is much less than even Mercury, at 5.5 percent of the mass of the Earth. Pluto's mass was too small to have produced deviations in Neptune's orbit, which were then found to be negligible.

Then, in 1992 after a five-year search using digital cameras and computerized analysis, David Jewitt and Jane Luu of the Massachusetts

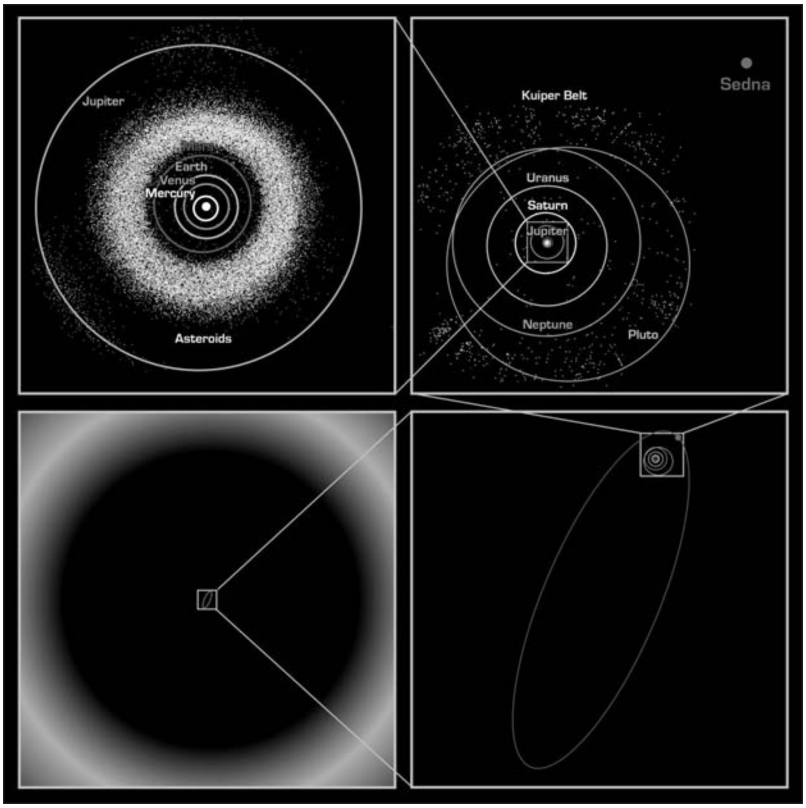
Institute of Technology (MIT) discovered the first of many similar icy objects beyond Neptune in a region called the Kuiper Belt. Existence of such a region had been predicted by Dutch American astronomer Gerard Kuiper. It is similar to the asteroid belt between Mars and Jupiter but about twenty times wider and populated by icy objects rather than the rocky and metallic bodies found in the asteroid belt. The Kuiper Belt extends from the orbit of Neptune, between 30 and about 55 astronomical units (AU), and is believed to contain thousands of objects larger than 100 kilometers in diameter.

More than 130 Kuiper Belt objects (KBOs) have been found with nearly the same 248-year period as Pluto at about 40 AU from the Sun. These "plutinos" complete their orbits twice during three orbits of Neptune, referred to as a 2:3 gravitational resonance. KBOs with other resonances, such as 3:5 and 4:7, are called cubewanos, and a few objects are found beyond a 1:2 resonance at 55 AU and with 330-year periods. Some objects have been found beyond 55 AU but are believed to have been scattered from the Kuiper Belt into a region called the scattered disk containing scattered disk objects (SDOs). Planetesimal objects in these latter two regions (KBOs and SDOs) are called trans-Neptunian objects (TNOs).

Astronomers began to view Pluto as the largest member of the new class of plutinos, and some started to question its status as a planet.



The dwarf planets Sedna and Pluto shown in size comparisons with other bodies. (NASA/JPL-Caltech/R. Hurt, SSC-Caltech)



Four panels show the location of the dwarf planet Sedna. (NASA/JPL-Caltech/R. Hurt, SSC-Caltech)

In 2003 a team from the California Institute of Technology (CalTech), working at Mount Palomar Observatory north of San Diego and led by Mike Brown, discovered an SDO at about 97 AU from the Sun, now called Eris. When a satellite was discovered in 2005, the mass of Eris was found to be 27 percent larger than that of Pluto, and a few astronomers began to refer to it as the tenth planet. Most astronomers, however, recognized that many TNOs might be larger than Pluto and that either they would also have to be classified as planets or Pluto would have to be reclassified to distinguish such objects from the traditional planets.

The definition of a planet was placed on the agenda for the General Assembly of the International Astronomical Union (IAU) meeting in August of 2006. An initial draft proposal recommended that Pluto be retained as a planet and that Ceres, Charon, and Eris be added to the list of planets. This recommendation was made by astronomers who viewed both Pluto and its satellite Charon as planets in a double-planet system, since each body rotates about a point located between the two. After many objections, an alternate proposal was offered by the Uruguayan astronomer Julio Fernández, who suggested an intermediate category for objects like Pluto, which are large enough to be nearly round but too small to clear their orbits of other planetesimals. The IAU accepted this proposal, and by unanimous vote it was agreed to call these intermediate objects "dwarf planets," with smaller objects to be called "small solar-system bodies." By further vote, it was agreed that Pluto is a dwarf planet.

These definitions as voted in Resolution 5A by the Twenty-sixth General Assembly of the IAU are as follows:

> The IAU therefore resolves that planets and other bodies in our Solar System, except satellites, be defined into three distinct categories in the following way:
>
> (1) A "planet" is a celestial body that (a) is in orbit around the Sun (b) has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape, and (c) has cleared the neighbourhood around its orbit.
>
> (2) A "dwarf planet" is a celestial body that (a) is in orbit around the Sun (b) has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape (c) has not cleared the neighbourhood around its orbit, and (d) is not a satellite.
>
> (3) All other objects, except satellites, orbiting the Sun shall be referred to collectively as "small solar-system bodies."

In three footnotes, this IAU resolution agreed that the eight planets are Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, and Neptune. It also agreed to establish a process for assigning objects to the category dwarf planet or another status. It also suggested that small solar-system bodies include most solar system asteroids, most TNOs, comets, and other small bodies. In the same meeting, the IAU announced only three members of the dwarf planet category: Ceres, Pluto, and Eris.

KNOWLEDGE GAINED

The new definitions of planets, dwarf planets, and small solar-system bodies have helped clarify both the nature of these objects and the structure of the solar system, as well as stimulating new research about them. The new definitions have led to new searches for dwarf planets and new research on criteria for hydrostatic equilibrium shape (nearly round) and orbital dominance (clearing the neighborhood).

The IAU maintains a dwarf planet watch-list of about a dozen candidates, which keeps changing as new candidates are found and as more is learned about the physics of existing candidates. Current candidates include the plutinos Orcus and Ixion, cubewanos Quaoar and Varuna, and the SDO Sedna, all of which are similar in size to or larger than Ceres (975 kilometers in diameter) but are not yet established as round. Observations indicate that icy bodies of more than about 400 kilometers reach hydrostatic equilibrium, but rocky objects with more rigid interiors might require at least 800 kilometers. The only other asteroid candidate seems to be Vesta, the second largest at 530 kilometers, which appears to be round except for a large impact crater. The Dawn space probe, scheduled to orbit Vesta by 2011, may resolve its status. Estimates range from forty to two hundred candidates in the Kuiper Belt and more beyond it.

CONTEXT

The new definitions of "planet" and "dwarf planet" highlight the increasingly complex nature of the solar system as more is discovered about it. The definitions have also, however, introduced many ambiguities and criticisms.

The new definitions do incorporate accepted theories for the evolution of the solar system and appeal to observational criteria. As planets formed from the dust and planetesimals of the solar disk, their gravity attracted more matter and they eventually dominated their orbits. However, if planetesimals were sufficiently disturbed by gravitational forces, such as those from nearby Jupiter, they never formed planets and remained as asteroids. Although no planets have completely cleared their orbital neighborhoods, even Mars, as the least dominant planet, has collected more than five thousand times as much material as that which remained in its orbit, while Ceres is only 0.33 times larger and Pluto only 0.07 times larger than the remaining material in their orbits.

Critics complained, however, that the new definitions were arbitrary, since no planet has completely cleared its orbit, and that the round shape of hydrostatic equilibrium is ambiguous, since there are various degrees of roundness. Others voiced concerns about the demotion of Pluto from its longtime status as a planet. Although the National Aeronautics and Space Administration (NASA) decided to accept the new definitions, many respected astronomers, including the director of the New Horizon mission to Pluto, Alan Stern, remained opposed, and his team continued to refer to Pluto as a planet. The discussions and debates would continue at later meetings of the IAU and as more was learned about solar-system objects and their physics.

Joseph L. Spradley

FURTHER READING

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory text covering the field of astronomy. Contains descriptions of astrophysical questions and their relationships.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An excellent introductory college text on planetary science by one of the leaders in the field. It has good chapters on the formation of the solar system and on asteroids and other small solar-system bodies.

Serge, Brunier. *Solar System Voyage*. Translated by Storm Dunlop. New York: Cambridge University Press, 2000. This well-illustrated book describes the solar system and discusses issues related to the definition of planets.

Sobel, Dava. *The Planets*. New York: Viking, 2005. A very readable account by a popular science writer of the nature and history of planets and asteroids, and of the scientists who study them.

Soter, Steven. "What Is a Planet?" *Scientific American* 132, no. 6 (January, 2007): 2513-2519. A planetary scientist discusses the controversy over the revised definition of a planet, including both its flaws and the scientific advantages of the concept of a dwarf planet.

Weintraub, David A. *Is Pluto a Planet? A Historical Journey Through the Solar System*. Princeton, N.J.: Princeton University Press, 2006. This book traces the concept of a planet from antiquity to the present day, providing the historical and astronomical context for deciding if Pluto is a planet.

See also: Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Eris and Dysnomia; Kuiper Belt; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Pluto and Charon.

E

Earth-Moon Relations

Categories: Earth; Planets and Planetology

The Moon is the closest astronomical body to the Earth, with a mass approximately 1.2 percent that of Earth. This unusually large fraction gives the Moon significant influence over the orbital and rotational motion of Earth, creating tides strong enough to have important geologic and oceanographic effects, among them variations in the length of the day. The Moon, along with the Sun, causes Earth's spin axis to precess with a period of 26,000 years.

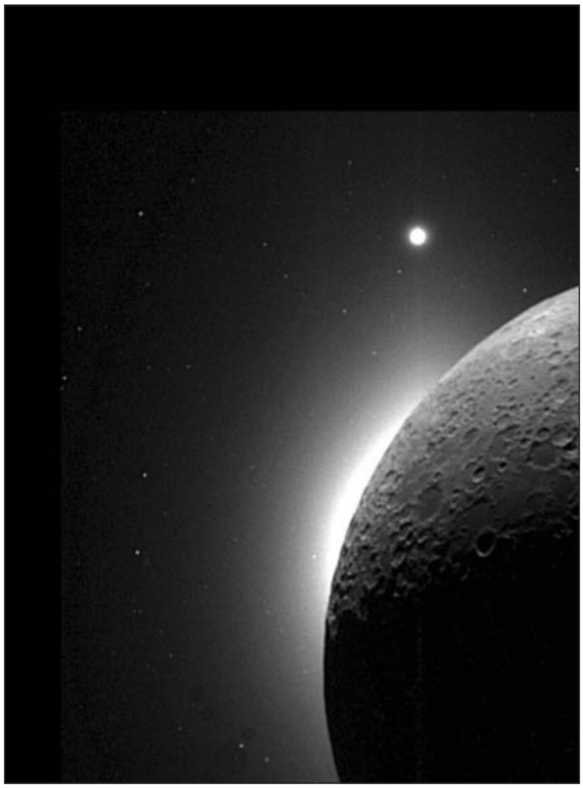
OVERVIEW

The Moon is the most prominent astronomical body after the Sun. It is the closest astronomical body to the Earth, orbiting at an average center-to-center separation of 384,000 kilometers. The Moon has a radius of 1,740 kilometers; at this distance, it appears to be 0.5° in angular width. The mass of the Moon is 3.74×10^{22} kilograms and the density of the Moon is 3.3 grams per cubic centimeter. Earth, by contrast, has a mass of 5.97×10^{24} kilograms and a radius of 6,380 kilometers, giving it a density of 5.5 grams per cubic centimeter, substantially more than that of the Moon. The lower density of the Moon, along with its lack of a magnetic field, argues that the Moon lacks a molten metallic core such as the Earth has.

Earth is close enough for material thrown off of the Moon by meteorite impact (called "ejecta") to fall onto it. A small number of meteorites discovered in desert areas or in Antarctica closely resemble lunar rocks collected by the Apollo astronauts and have been verified as of lunar origin.

The Moon and Earth are gravitationally bound. They orbit around a common point, called the barycenter, with a period of 27.3 days. This period is called the sidereal month and represents the time for the Earth-Moon system to complete one rotation with respect to the stars. The synodic month, by contrast, is 29.5 days, the time between successive full Moons.

The Earth-Moon system is gravitationally bound to the Sun. Hence, the barycenter orbits the Sun in obedience to Johannes Kepler's three laws of planetary motion: the orbit of the barycenter is an ellipse with the Sun at one focus; the line from the center of the Sun to the barycenter sweeps out equal areas in equal times; and, the cube of the radius of the barycenter orbit is proportional to the square of the period. The barycenter lies on a line joining the center of Earth to the center of the Moon, at



The Clementine spacecraft captured this image of the Moon in earthshine. The Sun peeks from behind the Moon, with Venus in the far background as the pearl of light near the top. (NASA) a point 4,680 kilometers from the center of Earth. This distance is 73 percent of the radius of Earth. An observer on Mars would see Earth displaced from the ideal elliptical orbit by as much as ⅜ of its diameter.

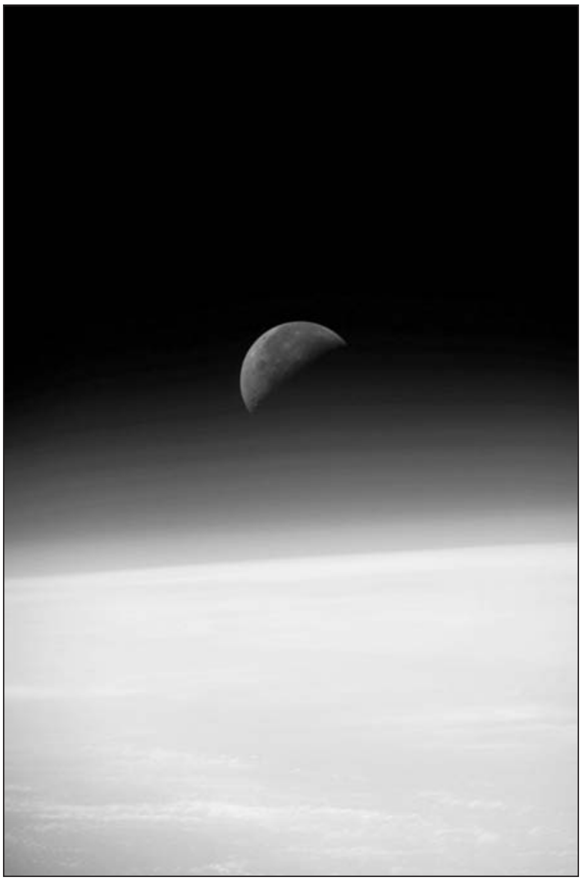
The motion of Earth about the barycenter is superimposed on the elliptical motion of the barycenter about the Sun in a complicated manner. Earth oscillates back and forth across the barycenter ellipse, spending half of a synodic month inside the ellipse (toward the Sun) and the other half outside the ellipse. Simultaneously, the Earth oscillates above and below the ecliptic (the plane of Earth's orbit) spending half of a sidereal month above the plane and the other half below it. These back-and-forth and up-and-down oscillations are not necessarily synchronous. When the Earth is inside the ellipse, the Moon is outside it, and vice versa. A similar arrangement holds for the up-and-down displacements. Absent the Moon, the center of the Earth would coincide with the barycenter and the planetary motion of the Earth would be close to the elliptical ideal. The Earth-Moon system, on the other hand, has one of the most complicated motions in the solar system.

The Galileo spacecraft returned separate images of Earth and its moon that were later compiled into this montage. (NASA/JPL)

The Earth-Moon system is like an unbalanced dumbbell tumbling end over end about the barycenter. The gravitational pull is the "bar" holding the dumbbell together. The sides of the Earth and Moon facing each other can be referred to as the "inner" sides, and the opposite sides of each can be referred to as the "outer" sides. The gravitational force falls off with distance, making the gravitational pull on the inner side of each body stronger than the gravitational pull on the outer side. This inequality of forces is referred to as gravitational tidal force. Neither the Earth nor the Moon is rigid. Each is plastic enough to change shape under the influence of the tidal force. The gravitational pull of the Moon raises a bulge in the Earth more or less directly under the Moon; the bulge is matched by a similar one at a location more or less directly opposite the Moon. The bulge in the ocean presents itself as the familiar tides. Similar but less familiar tides exist in the atmosphere and in the Earth's crust. The rotation of the Earth attempts to carry these bulges away from the point directly under the Moon, resulting in a slight sideways component to the mutual gravitational pull. This sideways pull acts as a brake on the rotational motion of the Earth, slowing it down and increasing the length of the day. The increase is approximately one-thousandth of a second per century, but it has been accumulating since the creation of the Moon billions of years ago. Growth-ring counts in fossil coral from 400 million years ago seem to indicate that the year (whose length should not change) consisted of about 400 days back then; now a year consists of about 365 days. In other words, the length of the day has increased by about 10 percent in the past 400 million years.

The sideways pull on the Moon is in the direction of its orbital motion around Earth. Extra energy imparted by the pull increases the radius of the Moon's orbit and also increases the length of the siderial month. Since the length of the day is increasing faster than the length of the sidereal month, eventually the two will become equal and the day and month will be the same. At that time, the Earth will always present the same face to the Moon, just as the Moon always presents the same face to the Earth today.

The rotation of the Earth gives it an oblate shape that is thicker at the equator than through the poles. The gravitational pull of the Sun and Moon on this equatorial bulge acts as a torque that causes the Earth to precess like a top. The spin axis of the Earth currently points toward Polaris, the pole star, but this is only an accident of history. In 13,000 years, Vega (in the constellation Lyra) will be the pole star.



The Moon viewed from Earth orbit, with the Earth's atmosphere rising above the planet's surface in the foreground. (NASA)

KNOWLEDGE GAINED

The bulk of Earth-Moon interactions are gravitational and are known from earthbound observations. The apparent location of the Sun in the zodiac on the first day of spring (recognized as the day that the Sun rose due East and set due West) held great cultural and religious significance to ancient civilizations and was monitored closely. Over the centuries, it became clear that this location, originally in the constellation Taurus, had moved to the constellation Aries. The Greek astronomer Hipparchus discovered this fact about 130 B.C.E. and from it deduced the 26,700-year circular motion of the north celestial pole. In 1530, Nicolaus Copernicus recognized this as due to drift of the Earth's rotational axis with respect to the fixed stars, and Sir Isaac Newton in 1687 showed the phenomenon to be an effect of Moon's gravitational influence on the Earth.

Edmond Halley in 1693 and Immanuel Kant in 1754 used Newtonian gravitational theory to calculate the locations, dates, and times of total solar eclipses discussed in ancient Greek and Roman documents. Their calculations argued that the eclipses could not have taken place at the dates and places recorded. The discrepancies were eventually traced to changes in the length of the day due to tidal braking.

Starting with Apollo 11, each subsequent lunar landing mission (except the ill-fated Apollo 13) brought back significant amounts of lunar rock for scientific study. Oxygen derived from the lunar material proved to have the same ratio of isotopes as oxygen found on Earth. In contrast, oxygen retrieved from meteorites believed to be of Martian origin had substantially different isotopic ratios.

This discovery, in conjunction with the observation that the Moon lacks an iron core, led to the impact theory of lunar origin. In this theory, the young Earth and a body approximately the size of Mars collided some 4.5 billion years ago. The collision threw a substantial amount of the

The Earth viewed from the Moon, as the Apollo 8 astronauts began their orbit on December 29, 1968. (NASA)

Earth's crust into space, where some of the material coalesced into the Moon, with the remainder falling back to Earth. Since this happened after the bulk of the iron in the proto-Earth had sunk into the core, the material that formed the Moon was relatively iron-free.

CONTEXT

The combined motion of the Earth and the Moon around their common barycenter is one of the most complicated problems in celestial mechanics. Newton once referred to it as the only problem that ever gave him a headache. Several factors complicate the solution. The influence of the Sun makes the problem a three-body gravitational interaction rather than the simpler two-body problem conquered by Kepler. Unlike the two-body problem, the three-body problem cannot be solved in closed analytic form; particular approximate solutions exist for special configurations, but the Sun-Earth-Moon trio does not conform to any of them.

The Earth and Moon are also too close for either to be regarded as point masses. Further, neither is purely spherical: the Earth is ellipsoidal, with an equatorial bulge as a product of its rotation; the Moon is oval as a result of a permanent tidal bulge on the side facing the Earth. The rotational period of the Moon equals its orbital period, so that one face perpetually faces the Earth, but the orbit is not circular, so that the Moon moves along the orbit at a varying rate. This causes the side of the Moon facing the Earth to rock back and forth, a motion known as libration. The deviation from circularity (called the eccentricity) is itself variable, driven by the gravitational pull of the Sun, so that the the extent of the libration waxes and wanes. This variation in eccentricity is called evection.

Billy R. Smith, Jr.

FURTHER READING

Brusche, P., and J. Sundermann. *Tidal Friction and the Earth's Rotation*. Berlin: Springer, 1978. A scholarly work. The first paper, "Pre-Telescopic Astronomical Observations" by F. R. Stephenson covers the historical eclipse data that revealed the slow increase in the length of the day due to tidal braking.

Comins, N. *What if the Moon Didn't Exist? Voyages to Earths That Might Have Been*. New York: HarperCollins, 1993. An astronomer examines how the Earth might have evolved without the interaction of a massive nearby Moon. An unusual but entertaining and engaging exploration of Earth-Moon interactions. Index and bibliography.

Darwin, G. *The Tides and Kindred Phenomena in the Solar System*. San Francisco: W. H. Freeman, 1962. A thorough discussion of tides from all perspectives: oceanographic, hydrodynamic, geological, astronomical, and historical. Contains a chapter on the less familiar tides in the atmosphere and in the Earth's crust. The remarks on tidal coupling between the Sun and planets are unfortunately out of date.

Ferguson, Kitty. *Tycho and Kepler: The Unlikely Partnership That Forever Changed Our Understanding of the Heavens*. New York: Walker, 2002. This book engagingly describes what is probably the most fruitful and important collaboration in all of the physical sciences. Tycho Brahe's naked-eye observations were the most accurate ever obtained before the invention of the telescope. They revealed hitherto unknown variations in the motion of the Moon; Brahe's observations of Mars made it possible for Kepler to discover the laws of planetary motion that bear his name. Well illustrated (includes six pages of color plates), with notes, bibliography, and index.

Kolerstrom, Nicholas. *Newton's Forgotten Lunar Theory: His Contribution to the Quest for Longitude*. Santa Fe, N.Mex.: Green Lion Press, 2000. Before the invention of the ship's chronometer, navigators used the motion of the Moon to determine longitude. Doing so accurately requires a very accurate theory of the Moon's orbital motion. In Newton's time, this was an area of immensely important scientific research. This book outlines his efforts at solving this exceedingly difficult problem. Chapter 1 describes the motions of the Sun, Earth, and Moon.

Moore, Patrick. *On the Moon*. London: Cassell, 2001. Contains an excellent nontechnical discussion of lunar motion and tides.

See also: Auroras; Earth's Age; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Origin; Earth's Rotation; Earth's Shape; Earth's Structure; Lunar History; Planetary Orbits: Couplings and Resonances; Solar System: Element Distribution; Van Allen Radiation Belts.

Earth-Sun Relations

Categories: Earth; Planets and Planetology; The Sun

The fundamental Earth-Sun relationship, from which all others derive, is that the Earth rotates on its axis as it revolves around the Sun. The relationships between the Earth and the Sun determine the Earth's "heat budget" and control life on the planet. Earth motions also produce noticeable periodic changes in the apparent path of the Sun across the sky, perhaps most obvious in the seasonal change in directions of sunrise and sunset, and the length of time the Sun is above the horizon.

OVERVIEW

Earth-Sun relations are the dominant controls of life on Earth. The Sun is a star, and its electromagnetic radiation warms the Earth and supplies the energy that supports life on the planet. Earth-Sun relations determine the amount, duration, and distribution of solar radiation that is received by Earth. The Earth's rotation on its axis produces day and night, and its revolution around the Sun and the tilt of its rotational axis result in the seasons; these processes serve to distribute solar radiation over the Earth. Earth's atmosphere and oceans influence the reflection, absorption, and transfer of solar energy. The result of these interacting phenomena is a "heat budget" on Earth that is hospitable to life.

The Sun radiates electromagnetic energy from every part of its spherical surface. Earth, 150 million kilometers away, intercepts only a minute portion of the Sun's radiation, about one two-billionth. The small amount of the Sun's energy that strikes Earth is Earth's energizer. It sustains life on Earth and drives weather systems and oceanic circulation. Solar energy from the past has been preserved in the form of fossil fuels—coal, petroleum, and natural gas.

Perhaps the most remarkable aspect of Earth—remarkable because it is rare in our solar system—is its relatively narrow range of moderate temperatures. The adjectives "hot" and "cold" are frequently used in describing our weather. In relation to the temperatures that are found elsewhere in the solar system, Earth is always moderate, and the words "hot" and "cold" better describe conditions on the other planets. The mean temperature of Earth is about 15° Celsius (59° Fahrenheit, or 288 kelvins); the absolute extremes recorded anywhere on Earth are 58° Celsius (136° Fahrenheit, or 331 kelvins) in North Africa and –89° Celsius (–128° Fahrenheit, or 184 kelvins) in Antarctica. Few inhabitants of Earth will ever experience a temperature range of much more than 60° or 70° Celsius in a lifetime. Compare these temperatures with those of Earth's nearest neighbor, the Moon, where temperatures range from about 120 to –170° Celsius (393 to 103 kelvins) between the day- and nightsides. Earth's so-called sister planet, Venus, has a surface temperature of about 450° Celsius (723 kelvins). The outer planets of the solar system experience a permanent deep freeze, below –100° Celsius (137 kelvins).

The best demonstration of the moderate nature of Earth's temperature is the presence of the world's oceans. Water can exist in the liquid state only in the narrow temperature range of 0 to 100° Celsius (273 to 373 kelvins) at Earth's surface atmospheric pressure. Yet almost 98 percent of Earth's water remains in the liquid state. The polar ice caps contain 2 percent, and a minute portion is water vapor in the atmosphere at any time. Currently 71 percent of Earth's surface is covered by oceans of liquid water, and it has had large oceans for much of its existence as a planet.

The factors that cause Earth to experience such moderate temperatures are complex and interrelated. The Sun is the source of the energy, yet being the right distance from the Sun cannot be the sole cause of Earth's moderate temperature—witness the Moon. Rather, the explanation has to do with Earth's atmosphere, its oceans, and its motions relative to the Sun.

Earth's atmosphere moderates the planet's temperature during both daylight and darkness. During daylight, the atmosphere blocks excessive amounts of solar radiation from reaching Earth's surface and thus prevents overheating. During darkness, the atmosphere retards the escape of heat in the form of long-wave infrared energy back into space and thus prevents excessive overnight cooling.

The oceans, also, have a pronounced effect on the heat budget of Earth. Water has the highest specific heat of any common substance. That means more heat is needed to raise the temperature of water than to raise the temperature of most other materials. Summers and daylight periods are kept cooler by the water's ability to absorb great amounts of solar energy without the water's temperature being raised significantly. During winter and during darkness, the water slowly gives up large amounts of heat without significant cooling of the water. Thus, the oceans act as a huge temperature buffer and, along with the atmosphere, add a moderating effect to temperature extremes.

Another factor in moderating temperature variations is the rate of Earth's rotation on its axis—one rotation in twenty-four hours. Rotation causes places on Earth to be alternately turned toward and away from the Sun, as though it were on a rotisserie. The relatively rapid rotation prevents places on Earth from overheating or overcooling. If Earth rotated significantly more slowly, so one side were exposed to the Sun for a much longer time, the illuminated side would become considerably hotter, while the dark side would cool down considerably more. For example, the planet Mercury rotates on its axis once in 58.6 Earth days, and it revolves around the Sun in 88 days; in other words, it makes two orbits around the Sun in the same time it completes three rotations on its axis. As a result, a given spot on the planet's surface is exposed to sunlight for 88 Earth days, and then is in darkness for 88 days more. The resultant temperature extremes range from about 430 to –170° Celsius (703 to 103 kelvins). Similarly, a given spot on the surface of the Moon is exposed to sunlight for a little over two Earth weeks and then is in darkness for about two weeks more, resulting in temperatures that range from about 120 to minus 170° Celsius (393 to 103 kelvins) between the day and night sides.

One complete revolution, or orbit, of Earth around the Sun defines the time unit of one year. During a single revolution, Earth rotates on its axis 365.25 times; therefore, there are 365 days in most calendar years, with an extra day every fourth year (leap year). The orbit of Earth around the Sun is an ellipse, which lies in a plane called the ecliptic plane. The Sun is located at one of the two foci of the ellipse; thus the distance of Earth from the Sun varies during the year. The point on the orbit where Earth is closest to the Sun is called perihelion; it occurs on or about January 3 each year at an Earth-Sun distance of about 147 million kilometers. The point on the orbit farthest from from the Sun is called aphelion; it occurs on or about July 4 each year at an Earth-Sun distance of about 152 million kilometers.

This variation in Earth's distance from the Sun does alter the amount of solar radiation that is received by Earth, but it is not the cause of the seasons. Perihelion, when Earth is nearest to the Sun and seemingly when Earth would be the warmest, occurs during winter in the Northern Hemisphere, and aphelion occurs during the Northern Hemisphere's summer. Thus the distance variations are out of phase with the seasons in the Northern Hemisphere, but in phase with seasons in the Southern Hemisphere. In both cases, the distance variations modify seasonal temperatures but do not cause the seasons themselves.

The cause of the seasons is the fact that Earth's axis of rotation is tilted 23.5° from the perpendicular to the ecliptic plane, which is the plane of Earth's orbit around the Sun. The orientation in space of Earth's rotational axis changes only very slowly, so that during one year (one orbit around the Sun) it remains nearly constant in position. As Earth revolves around the Sun, the axis in the Northern Hemisphere is alternately tilted toward and away from the Sun. When Earth's North Pole is tilted toward the Sun, the Northern Hemisphere receives more solar radiation than does the Southern Hemisphere, resulting in summer in the Northern hemisphere and winter in the Southern Hemisphere. When Earth's North Pole is tilted away from the Sun, the opposite occurs.

At the point in the orbit when the North Pole is tilted most directly away from the Sun, the Sun is exactly overhead at noon at the Tropic of Capricorn (23.5° south latitude) on Earth, and the entire area south of the Antarctic Circle experiences continuous daylight. This position in the orbit and the moment of time when it occurs both are referred to as the December solstice, which occurs around December 21 each year. For the Northern Hemisphere, it is the "winter solstice," but for the Southern Hemisphere, it is the "summer solstice." Six months later, when the North Pole is tilted most directly toward the Sun, the Sun is exactly overhead at noon at the Tropic of Cancer (23.5° north latitude) on Earth, and the entire area north of the Arctic Circle experiences continuous daylight. This position in the orbit and the moment in time when it occurs both are known as the June solstice, which occurs about June 21 each year. Approximately halfway in between the two solstices are the two equinoxes, occurring about March 20 or 21 and September 22 or 23 each year. On the two equinoxes, the Sun is directly overhead at noon at the Equator (0° latitude). Both Northern and Southern Hemispheres receive equal solar radiation then. On the two equinoxes, the periods of daylight and darkness are equal all over the Earth (equinox means "equal night"), and the Sun rises due east and sets due west.

Between March and September, when the Sun is overhead as seen from north of the Equator, sunrise is north of east and sunset is north of west for all locations that experience sunrise and sunset on a particular date, both Northern and Southern Hemispheres. (The places that do not experience sunrise and sunset are those areas near the poles that are experiencing either continuous daylight or continuous darkness on that date.) Sunrise is farthest north of east and sunset is farthest north of west on the June solstice, after which they both begin a southward migration. Between September and March, when the Sun is overhead as seen from south of the Equator, sunrise is south of east and sunset is south of west for all places that experience sunrise and sunset on a particular date, both Northern and Southern Hemispheres. (Again, the places that do not experience sunrise and sunset are those areas near the poles that are experiencing either continuous daylight or continuous darkness on that date.) Sunrise is farthest south of east and sunset farthest south of west on the December solstice, after which they both begin a northward migration to repeat the pattern.

METHODS OF STUDY

The seasonal variations in the Sun's apparent daily motion across the sky were noted by many ancient cultures. Various stone structures built hundreds to thousands of years ago around the world—from Stonehenge on England's Salisbury Plain to Caracol in Mexico's Yucatán peninsula to the Bighorn Medicine Wheel high in Wyoming's Bighorn Mountains to Mystery Hill in southern New Hampshire—display alignments pointing toward the rising and setting points of the Sun on the solstices and equinoxes.

More recently it has been determined that the tilt and orientation of Earth's rotational axis and the eccentricity of Earth's elliptical orbit change slowly and cyclically with time. The tilt of Earth's rotational axis relative to a perpendicular to the ecliptic plane (the plane of Earth's orbit around the Sun) now is about 23.5°, but it varies between approximately 21.5 and 24.5° over a cycle of 41,000 years. A greater tilt results in more extreme summer and winter climates, while a smaller tilt means summers are not as hot and winters are not as cold.

Earth's rotational axis also slowly wobbles like that of a giant top, tracing out in space a double cone over a period of 26,000 years. This wobble is due to the gravitational pull of the Moon and the Sun on Earth's equatorial bulge. At the present time, we are closest to the Sun during northern winter and farthest away during northern summer, but as a result of precession, in 13,000 years we will be farthest during northern winters and closest during northern summers, making seasons more severe in the Northern Hemisphere but less severe in the Southern Hemisphere.

Finally, the shape of Earth's orbit around the Sun slowly alternates between being more nearly circular and slightly more elliptical over a period of about 100,000 years due to gravitational perturbations by the other planets. The interplay of all these changes alter slightly the solar radiation received in the Northern and Souther Hemispheres and hence their seasonal climate variations. The Serbian astrophysicist Milutin Milanković was the first to study the effects of these changes and link them to the multiple advances and retreats of large-scale continental glaciation in the Northern Hemisphere during the Pleistocene epoch (the last two million years of geologic time).

Many solar phenomena such as sunspots, prominences, and flares vary in number and frequency of occurrence over a period of about eleven years (the solar activity cycle). This is caused by changes in the Sun's magnetic field, which reverses direction with each solar cycle. Long-term studies of the Sun show that the Sun's activity level varies over timescales of hundreds of years, becoming more or less active. The changes in solar activity seem to be related to changes in Earth's climate, as recorded in old documents and preserved in the width of annual tree rings. Very little solar activity was observed during the 1600's and 1700's, a period known as the Maunder minimum. During this time, Europe and northeastern North America were colder (the so-called Little Ice Age) and western North America experienced prolonged droughts. In contrast, solar activity seems to have been unusually high from sometime in the

1000's until about 1250, a time known as the Medieval Optimum (also known as the medieval grand maximum) when the climate was warmer than it is today. This time marked the height of the Vikings' expansion, when they established colonies in Greenland and Newfoundland. The colonies were abandoned or died out when solar activity declined and the climate turned colder.

Over still longer timescales, theories of stellar evolution applied to the Sun indicate that it has slowly increased in brightness since it formed about 4.5 billion years ago and that it will continue to do so for several billion years more, until it begins to run out of hydrogen fuel in its interior. When that happens, the Sun will expand relatively rapidly to become a red giant star several hundred to more than a thousand times brighter than it is now. These changes will greatly increase Earth's temperature, eventually making it uninhabitable.

Accurate measurements of the length of the day show small, erratic changes in Earth's rotation period, but on average an Earth day is lengthening by about 0.001 second every century. This slowing of Earth's rotation is due to tidal friction. The gravitational effects of the Moon and to a smaller extent the Sun produce tides on Earth, which gradually retard Earth's rotation. To conserve angular momentum, the Moon's distance from Earth is slowly increasing. This has been confirmed by accurately measuring the out-and-back travel time of laser beams bounced off retroreflectors left on the Moon's surface by the Apollo Moon landing missions. These processes will continue until, in the distant future (at least billions of years), Earth's rotation will become tidally locked with the Moon's revolution around Earth, both taking about 47 of our present days. However, this is happening so slowly that the Sun probably will become a red giant first.

Evidence of changes in Earth's heat budget in the past (from the recent past to the ancient past) have come from many disciplines, including history, geology, paleontology, climatology, and astronomy. Satellite studies of Earth over the past several decades have opened many new ways to monitor present conditions and look for predictors of possible future changes. For example, satellite images of the oceans at various wavelengths of the electromagnetic spectrum are analyzed to detect any slight temperature changes over time that may portend changes in Earth's climate. Sensitive satellite-borne instruments measure the intensity of sunlight in remote areas. Satellite imagery also provides an accurate record-base for changes in areas of snow cover in polar regions.

CONTEXT

Life on Earth is profoundly dependent upon the relationships between Earth and the Sun. The temperature of Earth is set by a balance between Earth's absorption of electromagnetic energy from the Sun and the subsequent reradiation of that energy from Earth as heat back into space. Life on the planet is dependent on this balance and the moderate temperatures that result.

Rotation influences Earth like a rotisserie, turning the planet so as to expose all sides to the Sun during the twenty-four-hour day for a more even heat. The atmosphere protects Earth from overheating by day and from overcooling at night. Earth's "greenhouse effect" is a result of the atmosphere's ability to trap solar radiation as heat during the day and retard its escape back into space at night, when the Sun is not above the horizon. Earth's heat budget is a product of many factors, not all of which are fully understood. Intense research continues on possible causes and effects of changes in Earth's heat budget. Being able to predict future changes is of prime importance so we can either prepare for them or try to avert them.

John H. Corbet

FURTHER READING

Ahrens, C. Donald. *Meteorology Today: An Introduction to Weather, Climate, and the Environment*. 8th ed. Florence, Ky.: Brooks/Cole, 2006. This introductory college-level text on meteorology presents a thorough treatment of weather phenomena and explains the seasons and the effects of solar energy on the atmosphere. Written for students with little background in science or mathematics. Includes many illustrations.

Chaisson, Eric, and Steve McMillan. *Astronomy*

Today. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Part of one chapter deals with Earth motions and the seasons; part of another, with solar activity.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Has sections dealing with sky motions, the seasons, and solar activity.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook, thorough and well written. Includes sections on sky motions, the seasons, and solar activity.

Gabler, Robert E., Robert J. Sager, Sheila M. Brazier, and D. L. Wise. *Essentials of Physical Geography*. 8th ed. Florence, Ky.: Brooks/Cole, 2006. A general introductory-level text on physical geography. Covers rotation, revolution, solar energy, and the elements of weather and climate. Well illustrated. Suitable for the general reader.

Harrison, Lucia Carolyn. *Sun, Earth, Time, and Man*. Chicago: Rand McNally, 1960. This book is considered the classic reference for Earth-Sun relations. Although dated, it is an excellent source of information and offers a remarkably extensive coverage of Earth-Sun relations.

Ruddiman, William F. *Earth's Climate: Past and Future*. 2d ed. New York: W. H. Freeman, 2008. A college textbook, suitable for both introductory and upper-level undergraduate courses. Contains much material on climate change and its causes.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses, divided into many short sections on specific topics. Has several sections on the motions of Earth, equinoxes and solstices, and solar activity.

Strahler, Arthur N., and Alan H. Strahler. *Modern Physical Geography*. 4th ed. New York: John Wiley & Sons, 1992. In this general college-level text on physical geography, Earth-Sun relations are discussed within the context of the study of weather and climate. Diagrams are well employed to explain Earth's orbit, rotation, revolution, and axis tilt. Easy to read.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. Several chapters in this introductory text deal with the solar system. The nature of solar activity and the Earth's motions are explained. Well illustrated and accessible.

See also: Auroras; Earth-Moon Relations; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Rotation; Earth's Shape; Eclipses; Greenhouse Effect; Main Sequence Stars; Solar Corona; Solar Flares; Solar Infrared Emissions; Solar Magnetic Field; Solar Radiation; Solar Radio Emissions; Solar Ultraviolet Emissions; Van Allen Radiation Belts.

Earth System Science

Categories: Earth; Scientific Methods

Earth system science views the planet Earth as a dynamic, unified system of simultaneous, interacting forces. In particular, Earth system science focuses on achieving a better understanding of the effects of human interaction with the environment.

OVERVIEW

A new approach to studying the Earth's systems views them as a set of interacting forces all operating simultaneously rather than separate Earth science disciplines to be studied in isolation. This new and promising viewpoint came about as a result of a growing recognition of the interactive nature of Earth's forces, as exerting influences on one another, as opposed to the idea that these forces act independently. The Earth is a constantly changing world with dramatic tectonic activity, volcanism, mountain building, earthquakes, dynamic oceans, severe storms, and varying climatic patterns and atmospheric conditions. Scientists using this "systems approach" view the Earth as a unified whole, and instead of concentrating attention on one component at a time, they use total global observation methods (attempting to model Earth as a whole) together with numerical modeling.

The Earth systems science approach was first detailed by an Earth System Science Committee (ESSC) appointed by the Advisory Council of the National Aeronautics and Space Administration (NASA). In 1986, the committee completed a three-year study of research opportunities in Earth science and recommended that an integrated, global Earth observation and information system be adopted and in full operation by the mid-1990's. The committee's *Overview Report* was released on June 26 of that year. Requests for the findings of the committee from the National Oceanic and Atmospheric Administration (NOAA) and the National Science Foundation (NSF), along with other federal agencies, have drawn the agencies—especially NASA, NOAA, and NSF—into a scientific alliance. The committee's report outlined immediate needs in several wide-reaching areas: scientific understanding of the entire Earth as a system of interacting components; the ability to predict both natural and human-induced changes in the Earth system; strong, coordinated research and observational programs in NASA, NOAA, and NSF as the core of a major U.S. effort; long-term measurements, from space and from Earth's surface, to describe changes as they occur and as a basis for numerical modeling; modeling, research, and analysis programs to explain the functioning of individual Earth system processes and their interactions; a sequence of specialized space research missions focusing on Earth systems, including the Upper Atmosphere Research Satellite (UARS), the joint United States/France Ocean Topography Experiment (TOPEX/POSEIDON), the Geopotential Research Mission (GRM), and an Earth-observing system using polar-orbiting platforms planned as part of the U.S. Space Station complex combining NOAA and NASA instrumentation.

Earth system science utilizes new technologies in global observations, space science applications, computer innovations, and quantitative modeling. These new tools of advanced technology allow scientists to probe and learn about the interactions responsible for Earth evolution and global change. Examples of research made possible by new tools are the opportunity to include the effects of global atmospheric motions in models of ocean circulation; the study of volcanic activity as a link between convection in the Earth's mantle and worldwide atmospheric properties; and the tracing of the global carbon cycle through the many transformations of carbon by biological organisms, atmospheric chemical reactions, and the weathering of Earth's solid surface and soils. In addition, recent advances in these technologies have had the immediate practical effect of improving the quality of human life in areas such as weather prediction, agriculture, forestry, navigation, and ocean-resource management.

The goal of Earth system science is to obtain a scientific understanding of the entire Earth system on a global scale by describing how its component parts and their interactions have evolved, how they function, and how they may be expected to continue to evolve on all timescales. This evolution is influenced by human activities—for example, the depletion of the Earth's energy and mineral resources and the alteration of atmospheric chemical composition—that sometimes are easily identified. The overall long-range consequences of these human actions are difficult to predict; the changes do not occur quickly enough for immediate recognition and, indeed, often take decades to evolve fully. The challenge to Earth system science is to develop the capability to predict those changes that will occur in the twenty-first century, both naturally and in response to human activity. To meet this challenge, vigorous investigations are being undertaken that include global observations, information systems built to process global data, and existing numerical models that already are contributing to a detailed understanding of individual Earth components and interactions. Such programs require interdisciplinary research support and interagency cooperation.

Observations from space, the best vantage point from which to obtain the comprehensive global data required to discriminate among worldwide processes operating on both long and short timescales, are essential to the study of the Earth as a system. Rapid variations in atmospheric and ocean properties, and the global effects of volcanic eruptions, ocean circulations, and motions of the Earth's crustal plates are examples of such processes. The Space Science Board of the National Academy of Sciences recommended orbital observation as a major method of global study; the Earth System Science Committee accepted the recommendations and expanded on them. Of particular value are NASA and NOAA satellites already on station in orbit, such as the Laser Geodynamics Satellites, which employ laser ranging to measure motions and deformations of Earth's crustal plates. Weather satellites already have supplied a sizable fund of data about the atmosphere and oceans, facilitating a good start on numerical modeling of weather variations. Other programs that have yielded coordinated studies of specific Earth system processes include the Earth Radiation Budget Experiment (1984), the Laser Geodynamics Satellites (1976 and 1983), the Navy Remote Ocean Sensing System (1985), and the Upper Atmosphere Research Satellite (1982).

In order to implement the full measure of the Earth system science concept, advanced information systems are needed to process global data and allow analysis, interpretation, and quantitative modeling. Also required is the implementation of additional satellite observations that yield ocean color imaging, scanning radar altimeters for surface topography, and atmospheric monitors. In addition, vigorous programs of ground-level measurements are needed to complement, validate, and interpret the global observations from space. International cooperation is essential to the success of Earth system science; the development of management policies and mechanisms are required to encourage cooperation among agencies around the globe in order to ensure the coordination necessary for a truly worldwide study of the Earth. A number of major international research programs are already operating, such as the World Climate Research Program, sponsored by the International Council of Scientific Unions and the World Meteorological Organization. To accomplish the many objectives of these programs, the Earth System Science Committee recommends two specific goals in which the three major U.S. agencies—NASA, NOAA, and NSF—must work closely together. The first goal is to establish and develop the advanced information systems and management structures required by Earth system science as a cooperative venture, and the second is to build close cooperation in programs of basic research.

The Orbiting Carbon Observatory which was lost during launch on February 24, 2009, was designed to be the first spacecraft to study Earth's atmospheric carbon dioxide, a main cause of global warming. (NASA/JPL)

METHODS OF STUDY

The most significant tools for global observation are Earth-orbiting satellites that can precisely measure large areas of the Earth at one time. Meteorological satellites, for example, gather enormous amounts of data about temperature, weather patterns and forces, and atmospheric changes and components; instruments aboard these satellites can gather data on and monitor variations of climate and storm systems, adding to the growing fund of global information. These satellites are placed in geosynchronous orbit at an altitude of 35,000 kilometers over the equator; at that altitude their orbital period is the same as the Earth's rotation—one day—so they remain over the same spot on Earth and can continuously monitor the same region.

Earth observation satellites, working in the infrared band of the spectrum, allow scientists to gather imagery and information about volcanic activity, earthquakes, geological formations, mineral resources, and geographic changes to provide still another perspective on the Earth. Orbiting the Earth from pole to pole many times a day, they are able to make a record of large sections of the Earth in a twenty-four-hour period as the Earth rotates underneath the satellite's orbital path. Earth observation satellites also carry instruments that measure temperatures, record cloud cover, and monitor catastrophic changes.

Other satellites measure ocean dynamics such as the temperature of large sections of seas and oceans, wave action, ocean water content, and relationships between water and the land it touches. Special instruments aboard these satellites are designed to monitor ice conditions and snowfall at sea and watch for changes in polar regions.

Still other spacecraft carry radar-imaging devices to measure precise distances and relationships between land features. The International Space Station has, as one of its most important objectives, the function of a permanently orbiting platform on which both humans and unattended instruments can work over long periods of time to monitor Earth activities and topography. The space station will be able to contribute large amounts of data because it can function both as information gatherer and processor using advanced onboard automated equipment such as specialized computers.

Although much of the instrumentation for Earth system science will be space-borne, much of it also will have to be ground-based, at the sites where data need to be gathered: near volcanoes, earthquakes, hurricanes, tornadoes, and thunderstorms, for example. Such phenomena must be measured on the ground to determine their effects on other Earth-surface processes. Ground data can then be compared and synthesized with data gathered from space to offer a broader view.

One of the most valuable of tools is the computer, for the receipt, storage, retrieval, analysis, and supply of large quantities of information. Ground-based and space-borne computers work in conjunction with each other for the comparison and large-scale analysis of data, which can be networked to any place on Earth. Computers also are used to generate theoretical models of various kinds of processes. By feeding weather data from the past hundred years into a computer, for example, scientists can begin to construct long-term models of weather patterns and global changes in climate and precipitation. Another study method is the creation and management of global information systems into which is fed data from countries all over the world; all nations can retrieve data for their own research as well as input data to add to the ongoing process of worldwide data analysis.

CONTEXT

The new methodology of Earth system science offers the opportunity to study the Earth from a more integrated perspective and to raise public awareness of the human practices that are affecting the planet. It is important that citizens of the twenty-first century understand the forces and processes that can cause global changes because, individually and collectively, they are contributors to those changes. Human contributions include continued clear-cutting of vast forest areas, thus inviting massive deforestation (destruction of forests); removal of protective trees and underbrush from areas adjacent to desert areas, thus encouraging rampant desertification (the spread of desert conditions);

and pollution of the atmosphere and waterways. Over time, these practices can slowly deplete Earth's natural resources and upset the fragile balance of nature worldwide. Human activities can trigger events that could cause long-term environmental damage.

Thomas W. Becker

FURTHER READING

Asrar, Ghassem. *EOS: Science Strategy for the Earth Observing System*. New York: American Institute of Physics, 1994. Describes the Earth Observing System program: its investigations, capabilities, and educational activities. For undergraduates and science readers.

Earth System Science Committee. *An Integrated Global Earth Observation and Information System to Be in Full Operation by the Mid-1990's*. Boulder, Colo.: University Corporation for Atmospheric Research, 1986. Written by the key people who created the method, a basic, concise presentation of the Earth Observation and Information System.

Kump, Lee R., James Kasting, and Robert Crane. *The Earth System*. Upper Saddle River, N.J.: Prentice Hall, 2003. An introductory work for those new to the Earth system field. Addresses the carbon cycle and events in Earth's history that help explain current global changes. For a general audience.

MacKenzie, Fred T. *Our Changing Planet: An Introduction to Earth System Science and Global Environmental Change*. Upper Saddle River, N.J.: Prentice Hall, 2002. Aimed at nonscientists, this volume covers all areas of Earth system science, including global change associated with both natural and human sources.

Matthews, Samuel W. "This Changing Earth." *National Geographic* 143 (January, 1973): 1-37. One of the earliest articles to describe the Earth's dynamic processes in a language that the public could readily understand. Addresses plate tectonics and takes the reader on a historic tour of the development of modern Earth science. Superb diagrams and supportive photography.

National Aeronautics and Space Administration Advisory Council. *Earth System Science Overview*. Washington, D.C.: Government Printing Office, 1986. This fifty-page document details in easy-to-understand language all the intricate natural mechanisms at work on the planet. Describes the entire Earth system science concept and outlines how the discipline's tools and methods will be brought together to focus on a global data-gathering, archiving of information, and international cooperative efforts. For high school students and general readers.

Skinner, Brian J. *The Blue Planet: An Introduction to Earth System Science*. New York: John Wiley, 1995. Good introduction to the field of Earth system science. Contains a series of essays covering both methods of research and current progress, some written by leading scientists. For a general audience.

See also: Auroras; Earth-Moon Relations; Earth-Sun Relations; Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Rotation; Earth's Shape; Earth's Structure; Eclipses; Gravity Measurement; Greenhouse Effect; Van Allen Radiation Belts.

Earth's Age

Category: Earth

Determining the age of the Earth is one of the great achievements of science. Until the eighteenth century, it was generally believed that the Earth was several thousand years old, and all its features—including mountains and valleys—had been produced by catastrophes such as great floods and earthquakes. The new science of geology gradually showed that the Earth was billions of years old and it had taking its form through slow, uniform processes operating over long periods of time.

OVERVIEW

In the middle of the seventeenth century, Joseph Barber Lightfoot of Cambridge University in England penned the following words:

> Heaven and Earth, center and circumference, were made in the same instant of time, and clouds full of water, and man was created by the Trinity on the 26th of October, 4004 B.C.E., at 9 o'clock in the morning.

At the time that Lightfoot wrote those words, this statement expressed the most informed opinion on the age of the Earth. The year 4004 B.C.E. had been calculated by James Ussher, the Anglican archbishop of Armagh, Ireland, by adding up the ages of the patriarchs recorded in the Old Testament. This was the method that most scholars used to date the Earth, and much effort was expended analyzing the first few books of the Old Testament.

A little over a century later, a Scottish physician and gentleman farmer named James Hutton (1726-1797) suggested that there was a better way to determine the past history of the Earth than by poring over biblical genealogies. Hutton believed that processes currently operating in nature could be extrapolated back in time to shed light on the historical development of the Earth. This idea—that past processes are essentially the same as present processes—is called uniformitarianism. In 1785, he presented his new views on geology in a paper entitled "Theory of the Earth: Or, An Investigation of the Laws Observable in the Composition, Dissolution, and Restoration of Land upon the Globe." Uniformitarianism became the foundation of the newly developing science of geology. Charles Lyell (1797-1875), who was born in the year of Hutton's death, extended these new ideas and helped lay the foundation for what was becoming a powerful new science. A major argument was over the age of the Earth. Was it really millions or billions of years old, as indicated by new discoveries and theories, or was it only a few thousand years old, as everyone had previously believed?

According to current theories, the matter that makes up the Earth as well as everything else in the universe was originally created in the big bang about 13 to 14 billion years ago. Hydrogen, most helium, and trace amounts of lithium and beryllium formed in the immediate aftermath of the big bang, while the atoms of all the other elements were formed by nuclear fusion reactions in massive stars as they generated energy during their "lives" and then exploded as supernovae.

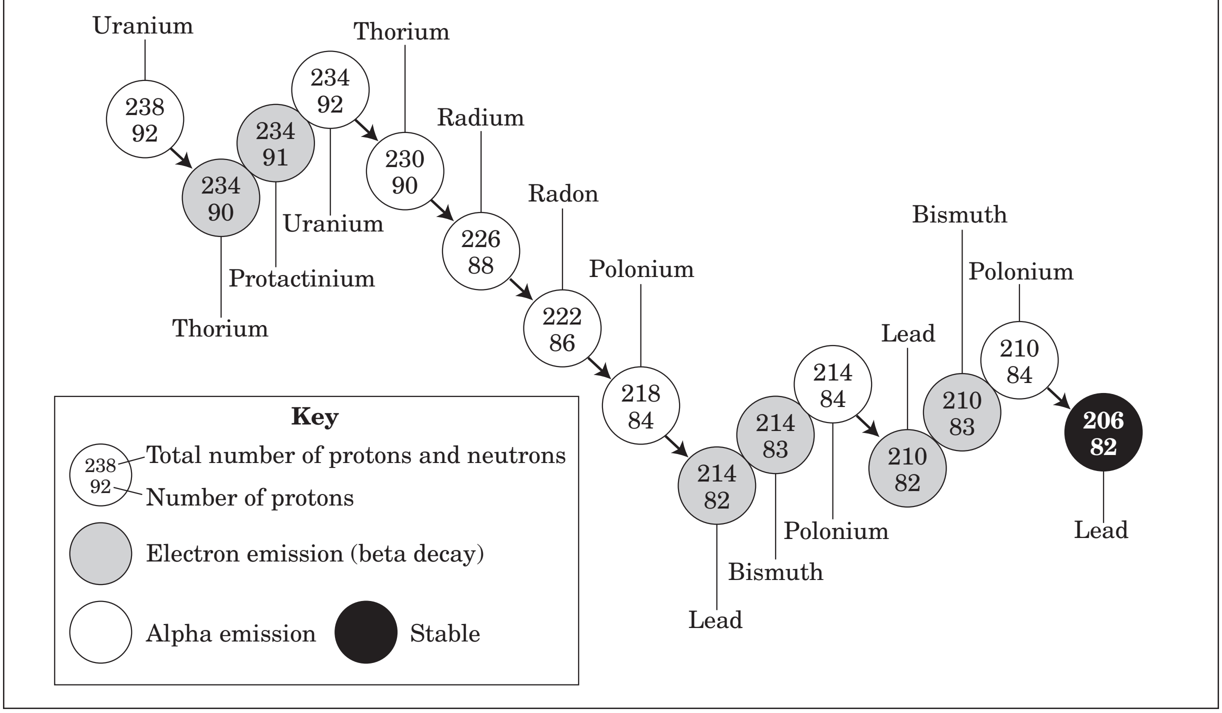
About 4.5 to 4.6 billion years ago, the matter that would become our solar system was part of a nebula, an interstellar cloud of gas and dust, in the disk of the Milky Way galaxy. The portion of this nebula that would become our solar system, called the solar nebula, began to contract under the influence of gravity. Most of the material in the solar nebula collapsed to the center and formed the Sun. The planets and everything else orbiting the Sun formed from the leftover debris through condensation and accretion. All this took place comparatively rapidly, over a time period of a few tens of millions to at most one hundred million years.

Since the initial formation of the Earth, many processes have been taking place in it and on it: Unstable atomic nuclei have radioactively decayed into nuclei of other elements; the Earth's early rotation rate has slowed due to friction from tides generated by the Moon and Sun; mountains have risen under the influence of global tectonics and have been worn away by the ceaseless activities of erosion; and evolution of life

Half-Lives of Some Unstable Isotopes Used in Dating

Parent Isotope	*Daughter Product*	*Half-Life Value*
Uranium 238	Lead 206	4.5 billion years
Uranium 235	Lead 207	704 million years
Thorium 232	Lead 208	14.0 billion years
Rubidium 87	Strontium 87	48.8 billion years
Potassium 40	Argon 40	1.25 billion years
Samarium 147	Neodymium 143	106 billion years

Source: U.S. Geological Survey.



Radioactive Decay of Uranium 238

has transformed the planet, changing barren wastelands into complex ecosystems teeming with diverse forms of life. These various processes have left their marks on the Earth; by studying them, scientists have begun to reconstruct the history of the Earth, in some cases all the way back to the origin of the Earth about 4.5 to 4.6 billion years ago.

METHODS OF STUDY

Many scientific attempts were made during the 1800's to try to determine the age of the Earth. Most involved some process that produced a noticeable change in something. By measuring the magnitude of the change and the rate at which it occurs, the age could be calculated. For example, suppose water is pouring into a bucket at the rate of 1 gallon per minute. If the bucket contains 10 gallons, one can calculate that the process started 10 minutes ago. However, the validity of that result hinges on several assumptions, such as (1) that the rate of water inflow has been constant, (2) that there was no water in the bucket at the start, and (3) that there are no holes in the bucket letting some water drain out.

Similar problems beset the early attempts at dating the Earth. Some of them included estimating how long it would take the Earth to cool and harden from a blob of molten rock and metal, how long it would take to accumulate the entire thickness of sedimentary rocks exposed all around the world, and how long it would take for the oceans to become as salty as they are. Because of uncertainty in the assumptions inherent in these methods, the calculated ages were not at all consistent, ranging from tens of thousands to hundreds of millions of years.

Much more accurate and consistent ages for Earth materials and events have been obtained by radiometric dating, a technique developed during the 1900's. This method uses radioactive decay, a process in which the unstable nucleus of an atom of one element (called the parent) spontaneously transforms into a nucleus of another element (called the daughter). Although it is impossible to say precisely when a single unstable parent nucleus will decay, a large number will decay at a definite rate referred to as half-life. A half-life is the period of time during which half the parent nuclei that are present will decay. Thus after one half-life, one-half of the parent nuclei will have decayed and one half will remain; after two half-lives, three-fourths will have decayed and one-fourth will remain; after three half-lives, seven-eighths will have decayed and one-eighth will remain; and so on. The decay rate or half-life appears to be constant, since nothing seems to have any effect on it. Measuring the ratio of daughter to parent in a sample can tell how long the parent has been decaying and thus how old the specimen is.

The nucleus of an atom is a dense packing of particles called protons with positive electrical charge and neutrons with no electrical charge. All nuclei of a particular chemical element have the same number of protons, but they can have different numbers of neutrons. Nuclei of the same element but with different numbers of neutrons are called isotopes of that element and are identified by their atomic mass number (which is simply the total number of protons and neutrons). For example, all carbon nuclei have six protons. Most carbon nuclei also have six neutrons, and this isotope is called carbon 12. Some carbon nuclei have eight neutrons, and that isotope is called carbon 14.

The three nuclear decay processes underlying geologic radiometric dating are (1) alpha decay, in which the parent nucleus emits an alpha particle, a helium nucleus of two protons and two neutrons; (2) beta decay, in which the parent nucleus emits a beta particle, an electron, when a neutron in the parent nucleus turns into a proton and electron; and (3) electron capture, in which the parent nucleus captures an electron that combines with a proton in the nucleus to form a neutron. The specific isotope decay schemes most commonly used for geologic dating are potassium 40 to argon 40 (half-life 1.3 billion years, via electron capture), rubidium 87 to strontium 87 (half-life 47 billion years, via beta decay), uranium 238 to lead 206 (half-life 4.5 billion years, via eight alpha and six beta decays), uranium 235 to lead 207 (half-life 713 million years, via seven alpha and four beta decays), and thorium 232 to lead 208 (half-life 14.1 billion years, via six alpha and four beta decays). Carbon 14 decays to nitrogen 14 via beta decay with a half-life of only 5,730 years; this is so short that it is limited to dating only very recent geologic events, although it has been very useful for archaeological and historical dating.

Radiometric dating has been applied to thousands of rock specimens from all over the Earth. The oldest rock formation found on Earth so far is the Acasta gneiss from near Great Slave Lake in northern Canada, dated at 4.03 billion years. Even older mineral grains—small crystals of zircon dated at 4.3 to 4.4 billion years—have been found in younger sedimentary rocks from the Jack Hills area of western Australia. It is difficult to find very ancient rocks on the surface of the Earth, because most of the Earth's surface has undergone many changes since the Earth was formed.

The currently accepted age for the Earth, 4.5 to 4.6 billion years, was obtained by radiometrically dating meteorites that fell to Earth from space. These meteorites are believed to be remnants from the time when the planets, including Earth, were forming in the early solar system. Similar dates have been obtained from a few of the rocks brought back by the Apollo landings on the Moon, which is believed to have formed at about the same time as the Earth.

CONTEXT

The problem of the age of the Earth is part of a much larger scientific question, which exists at the interface between the study of the Earth and its various processes (which often have practical benefits) and the more esoteric question of the origin and evolution of the universe as a whole. On the practical side, knowledge of the Earth is necessary to predict geologic disasters (such as earthquakes and volcanoes) and to search for geologic resources (such as oil and metallic ores). From a more esoteric point of view, the age of the Earth is important because it speaks to the most fundamental questions that are asked about our place in the universe: How old is this planet, and how was it formed? In the century or two since advances in geological science overthrew the seventeenth century notion of a much younger Earth, people have struggled with finding a new place in the universe. Proponents of "creation science" still argue that the Earth is thousands, not billions, of years old. Legal battles rage over the issue of whether schools across the United States should teach geochronology that is based on religious dogma rather than on scientific research. (In contrast, in Europe this is not a contentious issue, with little public questioning of the antiquity of the Earth based on geologic evidence.) While many questions remain about the details of the formation of the Earth, two facts seem clear: First, the Earth owes its origin to the same processes that brought the solar system into existence; second, those processes can be dated with a high degree of confidence at between 4.5 and 4.6 billion years ago.

Karl Giberson

FURTHER READING

Brush, Stephen G. *Nebulous Earth: The Origin of the Solar System and the Core of the Earth from Laplace to Jeffreys*. Cambridge, England: Cambridge University Press, 1997. Describes how thinking about the origin of the solar system changed and includes discussions of the origin of the Earth-Moon system.

Condie, Kent, and Robert Sloan. *Origin and Evolution of Earth: Principles of Historical Geology*. Upper Saddle River, N.J.: Prentice Hall, 1998. An easy-to-read text covering the complexities of the Earth’s history, life, and how Earth’s subsystems interact. Also discusses dating methods, planetary evolution, and ancient climates.

Dalrymple, G. Brent. *Ancient Earth, Ancient Skies: The Age of the Earth and Its Cosmic Surroundings*. Stanford, Calif.: Stanford University Press, 2004. A book designed for nonscientists who want to learn about the Earth’s age. Covers the manner in which scientists collect their data and describes the conclusions to which the data have led them.

Haber, Frances C. *The Age of the World: Moses to Darwin*. Baltimore: Johns Hopkins University Press, 1959. Reprint. Westport, Conn.: Greenwood Press, 1978. Focuses not on estimates of the age of the Earth but rather on the historical controversy that emerged when nonbiblical values for the age of the Earth began to be accepted. Provides insight into the conflict between science and dogma.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college textbook beyond the introductory level, with an approach based on comparative planetology. Offers much information on the origin of the solar system in general and the of Earth in particular.

Hurley, Patrick M. *How Old Is the Earth?* Garden City, N.Y.: Doubleday, 1959. One of the few full-length books on geochronology for the layperson. Even though published fifty years ago, it is still useful, as much of the broad overall outline has not changed appreciably since its publication, although many specific details have.

Lewis, Cherry. *The Dating Game: One Man’s Search for the Age of the Earth*. Cambridge, England: Cambridge University Press, 2002. The story of Arthur Holmes and the evolution of calculating the age of rocks. Written by a geologist who makes the more technical details understandable to a general audience.

Ozima, Minoru. *The Earth: Its Birth and Growth*. Cambridge, England: Cambridge University Press, 1981. A translation of a Japanese book that was written by a scientist whose specialty is geochronology. Written at an introductory level.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. It has two chapters on geologic dating and the historical development of the Earth.

Thackray, John. *The Age of the Earth*. New York: Cambridge University Press, 1989. About forty pages long, and published by a British geological museum, this concise volume contains more pictures than text, but the pictures, most in color, are helpful and make this an interesting source.

Wicander, Reed, and James Monroe. *Historical Geology*. 5th ed. Florence, Ky.: Brooks/Cole, 2006. An undergraduate text covering all major areas of historical geology. Provides a

history of the Earth and events that have shaped it.

See also: Earth-Moon Relations; Earth-Sun Relations; Earth System Science; Earth's Atmosphere; Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Rotation; Earth's Shape; Earth's Structure; Solar System: Origins.

Earth's Atmosphere

Category: Earth

The chemical composition of the atmosphere has changed significantly over the history of the Earth. The composition of the atmosphere has been influenced by a number of processes, including interaction with the solar wind; "outgassing" of volatiles (materials that easily vaporize to form gases) originally trapped in the Earth's interior during its formation; the geochemical cycling of carbon, nitrogen, hydrogen, and oxygen compounds between the surface, the ocean, and the atmosphere; and the origin and evolution of life.

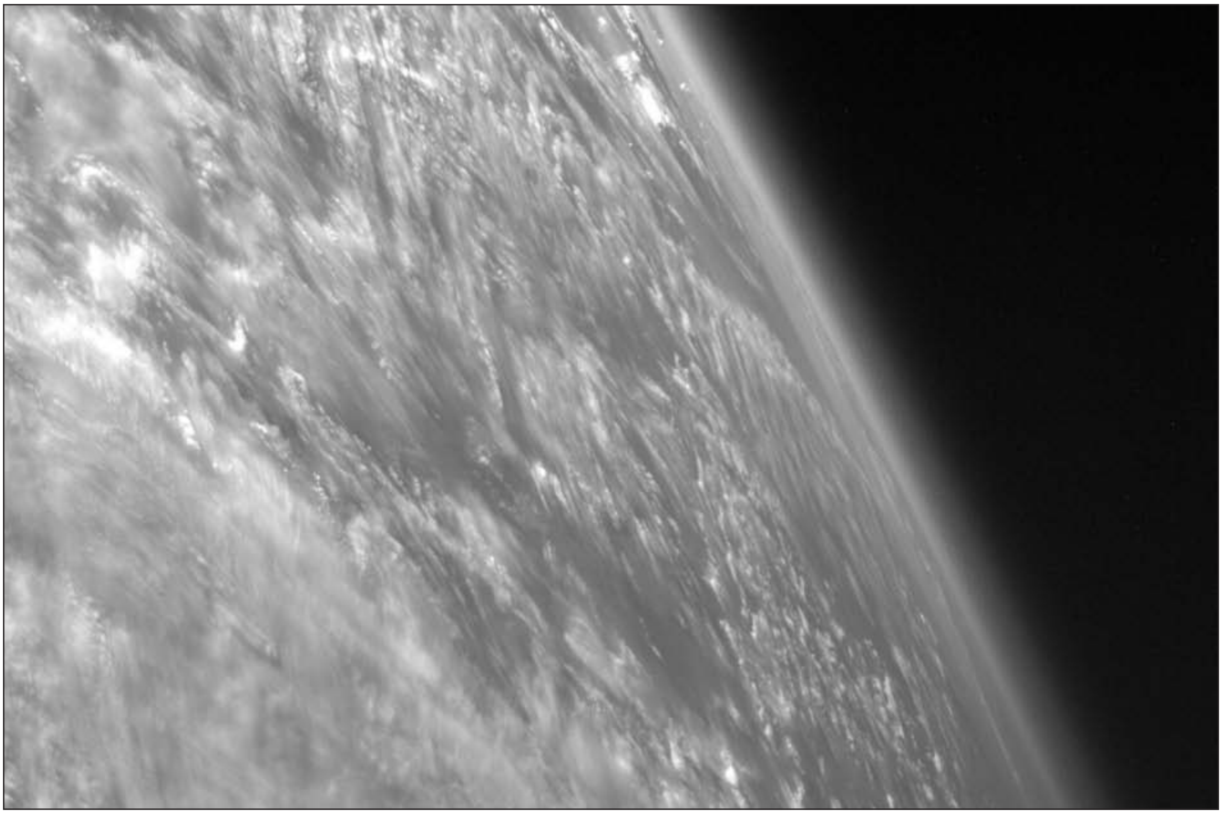
OVERVIEW

About 4.5 to 4.6 billion years ago, the primordial solar nebula, a part of a large interstellar cloud of gas and dust, began to contract under the influence of gravity. This contraction led to the formation of the Sun and the rest of the solar system including the Earth. The primordial solar nebula was composed mostly of hydrogen gas, with a smaller amount of helium, still smaller amounts of carbon, nitrogen, and oxygen, and still smaller amounts of the rest of the elements of the periodic table.

As the solar nebula contracted, its rotational speed increased to conserve angular momentum. Most of its mass contracted to its center, there becoming the proto-Sun, while the remaining matter was spun off into an equatorial disk. Within the disk, matter condensed from the gaseous state into small, solid grains. Only materials with high melting-point temperatures could condense near the developing proto-Sun; materials with lower melting points condensed farther out. The small solid grains collided with each other and stuck together in a process called accretion that led to the growth of larger bodies called planetesimals. Continued accretion resulted in fewer but larger planetesimals that eventually grew into protoplanets and finally planets, such as Earth.

About the time that the newly formed Earth was reaching its approximate present mass, it may have acquired a temporary atmosphere of hydrogen, helium, methane, ammonia, water vapor, and carbon dioxide—gases that were common in the solar nebula. However, even if such an atmosphere had surrounded the very young Earth, it would have been very short-lived. The Earth's mass was too small to have enough gravity to retain hydrogen and helium for long, and those gases would have quickly escaped into space. Ammonia and methane were chemically unstable in the early Earth's environment and were readily destroyed by ultraviolet radiation from the young Sun. Also, as the young Sun went through its T-Tauri phase of evolution, very strong solar winds (the supersonic flow of protons and electrons from the Sun) would have quickly stripped most of the rest of this primitive atmosphere away.

The early Earth was heated by the intense bombardment of remaining planetesimals and the decay of radioactive elements, to the point where it at least partly melted. The heating and melting released volatiles trapped in its interior by a process called outgassing, forming a gravitationally bound atmosphere. (It is believed that the atmospheres of Mars and Venus also originated in this manner.) The period of extensive volatile outgassing may have lasted for many tens of millions of years. The outgassed volatiles probably had roughly the same chemical composition as do present-day volcanic gaseous emissions: by volume about 80 percent water vapor, 10 percent carbon dioxide, 5 per-



This oblique photograph of Earth, taken from the space shuttle on September 4, 1997, shows the atmosphere and cloud cover over the northwestern African continent. (NASA/JPL/UCSD/JSC)

cent sulfur dioxide, 1 percent nitrogen, and smaller amounts of hydrogen, carbon monoxide, sulfur, chlorine, and argon.

The water vapor that outgassed from the interior soon reached its saturation point, which is controlled by the atmospheric temperature and pressure. Once the saturation point was reached, the atmosphere could not hold any additional gaseous water vapor. Any new outgassed water vapor that entered the atmosphere would have precipitated out of the atmosphere as rain that fell and formed the Earth's vast oceans. Only small amounts of water vapor remained in the atmosphere—ranging from a fraction of a percentage point to several percent by volume, depending on atmospheric temperature, season, and latitude.

The outgassed atmospheric carbon dioxide, being very water-soluble, readily dissolved into the newly formed oceans and formed carbonic acid. In the oceans, carbonic acid formed ions of hydrogen, bicarbonate, and carbonate. The carbonate ions reacted with ions of calcium and magnesium in the ocean water, forming carbonate rocks, which precipitated out of the ocean and accumulated as seafloor carbonate sediments. Most of the outgassed atmospheric carbon dioxide formed carbonates, leaving only trace amounts of gaseous carbon dioxide in the atmosphere (about 0.035 percent by volume).

Sulfur dioxide, the third most abundant component of volatile outgassing, was chemically transformed into other sulfur compounds and sulfates in the atmosphere. Eventually, the sulfates formed atmospheric aerosols and diffused out of the atmosphere onto the surface.

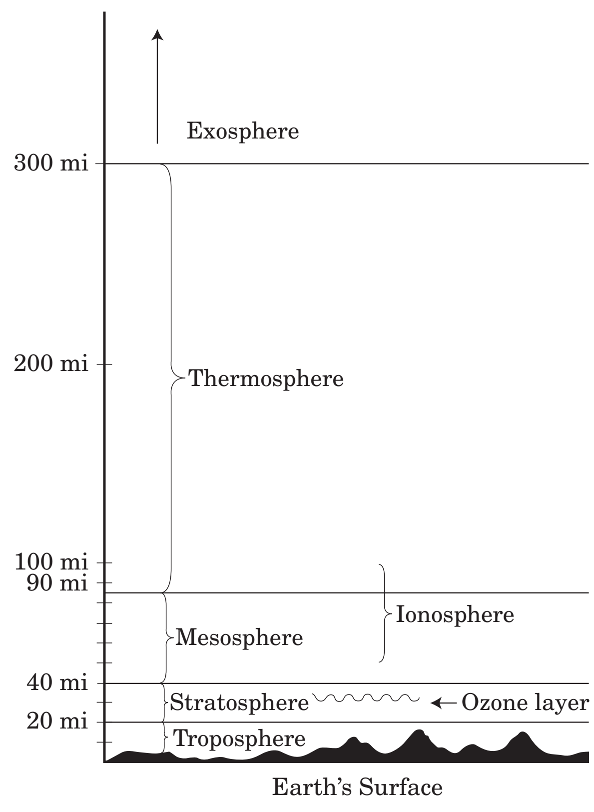
The fourth most abundant outgassed component, nitrogen, is chemically inert in the atmosphere and thus was not chemically transformed, as was sulfur dioxide. Unlike carbon dioxide, nitrogen is relatively insoluble in water and, unlike water vapor, does not condense out of the atmosphere. For these reasons, nitrogen remained in the atmosphere to become its major constituent (now 78.08 percent by volume). In this way, volatile outgassing led to the formation of the Earth's atmosphere, oceans, and carbonate rocks.

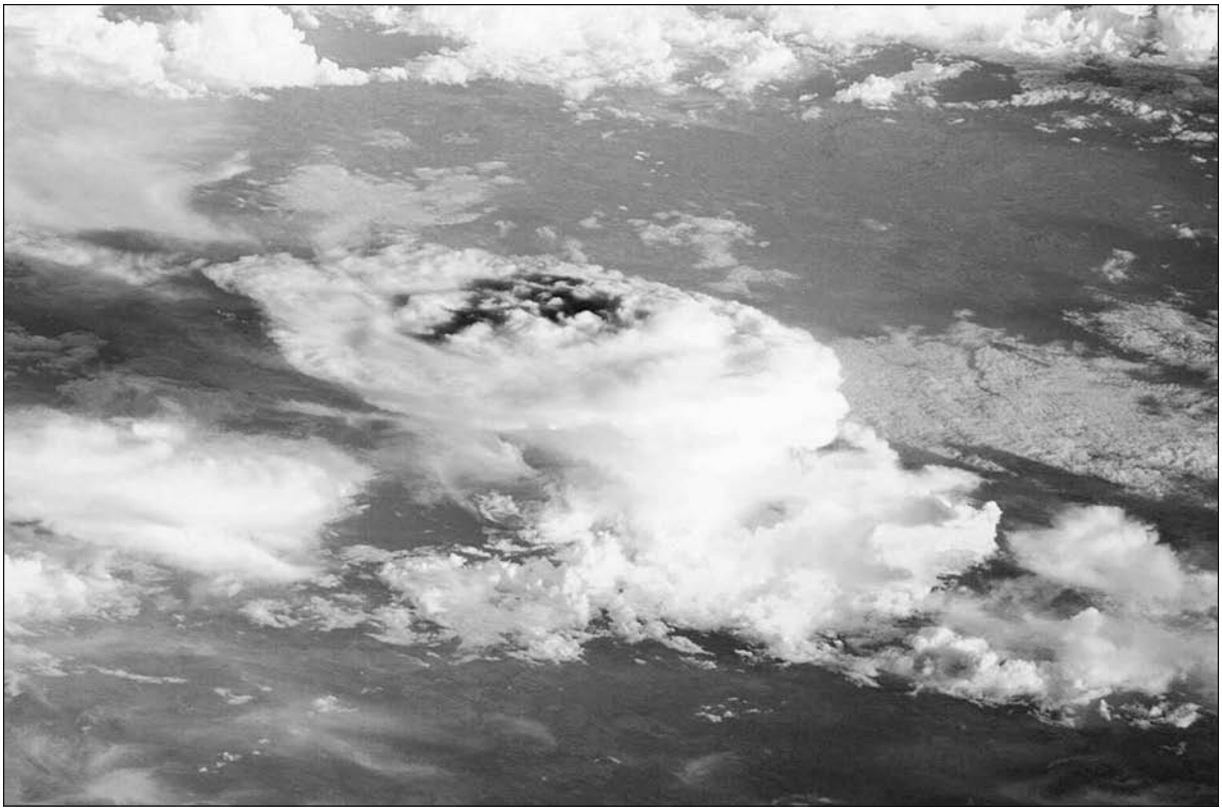
The molecules of nitrogen, carbon dioxide, and water vapor in the early atmosphere were acted upon by solar ultraviolet radiation and atmospheric lightning. In the process, molecules of formaldehyde and hydrogen cyanide could have been chemically synthesized, which would have precipitated and diffused out of the atmosphere into the oceans. In the oceans, the formaldehyde and hydrogen cyanide may have entered into polymerization reactions that eventually led to the chemical synthesis of amino acids, the building blocks of living systems. The synthesis of amino acids from nitrogen, carbon dioxide, and water vapor in the atmosphere and ocean is called chemical evolution. Chemical evolution preceded and provided the material for biological evolution.

There is chemical trace evidence for the existence of microbial living organisms on the Earth by about 3.8 billion years ago; the oldest known simple fossils are at least 3.5 billion years old. These earliest living organisms were anaerobic since there was no free oxygen in the atmosphere and oceans. Photosynthesis evolved in one or more of these early microbial groups, such as cyanobacteria. In photosynthesis, the organism utilizes water vapor and carbon dioxide in the presence of sunlight and chlorophyll to form carbohydrates, used by the organism for food. In the process of photosynthesis, oxygen is given off as a metabolic by-product. The production of oxygen by photosynthesis was a major event on the Earth and transformed the composition and chemistry of the early atmosphere. As a result of photosynthetic production, oxygen built up to become the second most abundant constituent of the atmosphere (now 20.9 percent by volume).

The evolution of atmospheric oxygen had very important implications for the evolution of life. The presence and buildup of oxygen led to the evolution of respiration, which replaced fermentation as the energy production mechanism in living systems. Accompanying and directly controlled by the buildup of atmospheric oxygen was the origin and evolution of atmospheric ozone, which is chemically formed from oxygen. The production of atmospheric ozone resulted in shielding the Earth's surface from biologically lethal solar ultraviolet wavelengths between about 200 and 300 nanometers. Prior to the evolution of the atmospheric ozone layer, early life was restricted to a depth of at least several meters below the ocean surface. At this depth, the ocean water offered shielding from solar ultraviolet radiation. The development of the atmospheric ozone

Layers of the Earth's Atmosphere



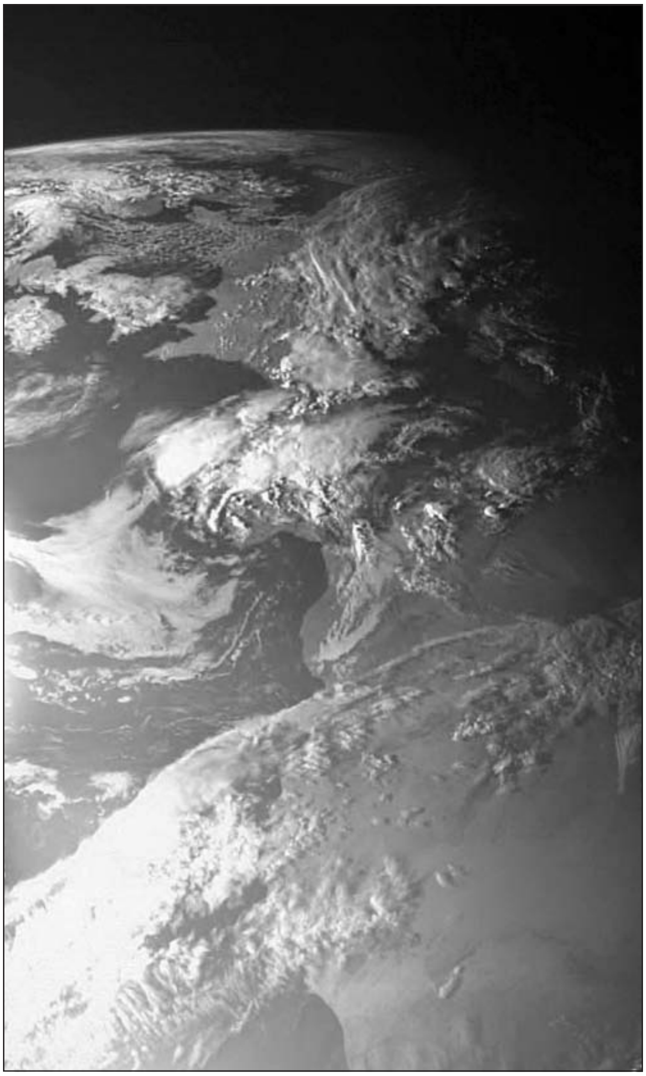


Storm clouds hover over Earth in this image taken from the International Space Station. (NASA)

layer and its consequent shielding of the Earth's surface permitted early life to leave the safety of the oceans and go ashore for the first time in the history of the planet. Theoretical computer calculations indicate that atmospheric ozone provided sufficient shielding from biologically lethal ultraviolet radiation for the colonization of the land once oxygen reached about one-tenth of its present atmospheric level.

Mercury, Venus, and Mars formed in a fashion similar to Earth, but they developed very differently because of their masses and distances from the Sun. They all experienced a period of heating, partial melting, and volatile outgassing of the same gases that led to the formation of the Earth's atmosphere. However, Mercury's distance from the Sun is so close (resulting in high temperatures) that its relatively weak gravity (due to its small mass) was unable to retain more than a thin trace of gases, and thus today it has virtually no atmosphere. In the case of Venus and Mars, the important difference is that the outgassed water vapor never existed in the form of liquid water on the surfaces of those two planets.

Because of Venus's closer distance to the Sun (108 million kilometers versus 150 million kilometers for Earth), its lower atmosphere was too hot to permit the outgassed water vapor to condense out of the atmosphere. Thus, the outgassed water vapor remained in gaseous form in the atmosphere and, over geological time, was broken apart by solar ultraviolet radiation to form hydrogen and oxygen. The very light hydrogen gas quickly escaped from the atmosphere of Venus, and the heavier oxygen combined with surface minerals to form a highly oxidized surface. In the absence of liquid water on the surface of Venus, the outgassed carbon dioxide remained in the atmosphere and built up to become the overwhelming constituent of the atmosphere of Venus (about 96 percent by volume). The outgassed nitrogen accumulated to comprise only about 4 percent by volume of the atmosphere of Venus. This carbon dioxide and nitrogen atmosphere is very massive—it produces an atmospheric pressure at the surface of the Venus about ninety times the surface pressure of Earth's atmosphere. If the outgassed carbon dioxide in the atmosphere of Earth had not left via dissolution in the oceans and resultant carbonate rock formation, the Earth's surface atmospheric pressure would be about seventy times greater than at present, with carbon dioxide comprising about 98-99 percent of the atmosphere and nitrogen about 1-2 percent. Thus, the atmosphere of Earth would closely resemble that of Venus. The thick carbon dioxide atmosphere of Venus causes a very significant greenhouse temperature enhancement, giving the lower atmosphere and surface of Venus a temperature of about 750 kelvins (about 477° Celsius), which is hot enough to melt lead. For comparison, the average surface temperature of Earth is only about 288 kelvins (about 15° Celsius).



The SMART 1 spacecraft took this image of Earth from 70,000 kilometers above the surface in May, 2004. Clouds and weather patterns are visible from Scandinavia (top) to northwestern Africa. (European Space Agency)

Like Venus, Mars has an atmosphere composed primarily of carbon dioxide (about 95 percent by volume) and nitrogen (about 3 percent by volume). Because of Mars's greater distance from the Sun (228 million kilometers versus 150 million kilometers for Earth), the temperature of the surface of Mars was too low to support the presence of liquid water. There may be very large quantities of outgassed water in the form of ice or permafrost below the surface of Mars. In the absence of liquid water, the outgassed carbon dioxide remained in the atmosphere. The atmospheric pressure at the surface of Mars, however, is only about 7 millibars (the average surface atmospheric pressure on Earth is 1,013 millibars). The smaller mass of the atmosphere of Mars compared to the atmosphere of Venus and Earth may be attributable to the smaller mass of Mars and, therefore, the smaller mass of volatiles trapped in the interior of Mars during its formation. In addition, it appears that the amount of gases trapped in the interiors of Venus, Earth, and Mars during their formation decreased with increasing distance from the Sun. Venus appears to have trapped the greatest amount of gases and was the most volatile-rich planet, Earth trapped the next greatest amount, and Mars trapped the smallest amount.

The atmospheres of the outer planets—Jupiter, Saturn, Uranus, and Neptune—all contain appreciable quantities of hydrogen and helium, along with methane and ammonia. It is believed that the atmospheres of these planets, unlike the atmospheres of the terrestrial planets Venus, Earth, and Mars, are captured remnants of the primordial solar nebula. Because of the outer planets' large masses and their great distance from the Sun resulting in their very low temperatures, hydrogen, helium, methane, and ammonia are stable and long-lived constituents of their atmospheres.

METHODS OF STUDY

Information about the origin, early history, and evolution of the Earth's atmosphere comes from a variety of sources. Information on the origin of Earth and other planets is based on theoretical computer simulations. These computer models simulate the collapse of the primordial solar nebula and the formation of the planets. Astronomical observations of what appear to be equatorial disks and the possible formation of planetary systems around young stars have provided new insights into the computer modeling of this phenomenon. Information about the origin, early history, and evolution of the atmosphere is based on theoretical computer models of volatile outgassing and the geochemical cycling and photochemistry of the outgassed volatiles. The process of chemical evolution—which led to the synthesis of organic molecules of increasing complexity, the precursors of the first living systems on the early Earth—is studied in laboratory experiments. In these experiments, mixtures of gases simulating the Earth's early atmosphere are energized by ultraviolet radiation, electrical discharges, or heated rocks, simulations of energy sources available on the early Earth. The resulting products are analyzed by chemical techniques.

One of the parameters affecting atmospheric photochemical reactions, chemical evolution, and the origin of life was the flux of solar ultraviolet radiation incident on the early Earth. Astronomical measurements of the ultraviolet emissions from young, sunlike stars have provided important information about ultraviolet emissions from the young Sun during the very early history of the atmosphere.

Geological and paleontological studies of the oldest rocks and the earliest fossil records have provided important information on the evolution of the atmosphere and the transition from an oxygen-deficient to an oxygen-sufficient atmosphere. Studies of the biogeochemical cycling of the elements have provided important insights into the later evolution of the atmosphere. Thus, studies of the origin and evolution of the atmosphere are based on a broad cross-section of the sciences, involving astronomy, geology, geochemistry, geophysics, and biology as well as atmospheric chemistry.

CONTEXT

Studies of the origin and evolution of the atmosphere have provided new insights into the processes and parameters responsible for global change. Understanding the history of the atmosphere provides insight into its future. Today, atmospheric changes being studied for their possible long-term effects include the buildup of greenhouse gases like carbon dioxide and the depletion of ozone in the stratosphere. The study of the evolution of the atmosphere has provided new insights into the biogeochemical cycling of elements between the atmosphere, biosphere, land, and ocean. Understanding this cycling is a key to understanding environmental problems and possible remedies. Studies of the origin and evolution of the atmosphere have also provided new insights into the origin of life and the possibility of life outside the Earth.

Joel S. Levine

FURTHER READING

Ahrens, C. Donald. *Essentials of Meteorology: An Invitation to the Atmosphere*. 5th ed. Florence, Ky.: Brooks/Cole, 2007. An updated version of a classic meteorology textbook. Explains tricky concepts in an easy-to-understand way. Suitable for students and nonscientists.

_______. *Meteorology Today*. 8th ed. Florence, Ky.: Brooks/Cole, 2006. A common text used for introductory meteorology college courses but can also be understood by general audiences. Comes with a CD-ROM learning aid, which includes chapter tests and multimedia tutorials.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses with an entire chapter on the formation of the planets.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A textbook for introductory astronomy courses that offers several sections dealing with the origin of the solar system and planetary atmospheres.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook with several sections on the origin of the solar system, including coverage of planetary atmospheres.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college textbook beyond the introductory level, its approach is based on comparative planetology. Includes information throughout the text on the origin of the solar system and its planetary atmospheres.

Henderson-Sellers, A. *The Origin and Evolution of Planetary Atmospheres*. Bristol, England: Adam Hilger, 1983. A technical treatment of the variation of the atmosphere of the Earth over geological time and the processes and parameters that controlled it. Chapters cover the mechanisms for long-term climate change, the atmospheres of the other planets, planetary climatology on shorter timescales, and the stability of planetary environments.

Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans*. Princeton, N.J.: Princeton University Press, 1984. A comprehensive and technical treatment of the geochemical cycling of elements over geologic time and the coupling between the atmosphere, ocean, and surface. Includes coverage of the origin of the solar system, the release and recycling of volatiles, the chemistry of the early atmosphere and ocean, the acid-base balance of the atmosphere-ocean-crust system, and carbonates and clays.

Levine, Joel S., ed. *The Photochemistry of Atmospheres: Earth, the Other Planets, and Comets*. Orlando, Fla.: Academic Press, 1985. A series of review papers dealing with the origin and evolution of the atmosphere, the origin of life, the atmospheres of Earth and other planets, and climate. The book contrasts the origin, evolution, composition, and chemistry of Earth's atmosphere with the atmospheres of the other planets. It contains two appendixes that summarize all atmospheric photochemical and chemical processes.

Lewis, John S., and Ronald G. Prinn. *Planets and Their Atmospheres: Origin and Evolution*. New York: Academic Press, 1983. A comprehensive treatment of the formation of the planets and their atmospheres. Begins with a detailed account of the origin and evolution of solid planets via coalescence and accretion in the primordial solar nebula; then discusses the surface geology and atmospheric composition of each planet.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A thorough college textbook for introductory astronomy courses, divided into many short sections on specific topics. Contains a unit on the origin of the solar system and several sections on the origins of planetary atmospheres.

Schopf, J. William, ed. *Earth's Earliest Biosphere: Its Origin and Evolution*. Princeton, N.J.: Princeton University Press, 1984. A comprehensive group of papers on such subjects as the early Earth, the oldest rocks, the origin of life, early life, and microfossils. Chapters address the oldest known rock record, prebiotic organic syntheses and the origin of life, Precambrian organic geochemistry, the transition from fermentation to anoxygenic photosynthesis, the development of an aerobic environment, and early microfossils. Technical.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and fully illustrated, with a chapter on the origin and historical development of the Earth and its atmosphere.

See also: Auroras; Earth-Sun Relations; Earth's Origin; Earth's Rotation; Earth's Shape; Earth's Structure; Eclipses; Greenhouse Effect; Planetary Atmospheres; Van Allen Radiation Belts.

Earth's Composition

Category: Earth

The Earth consists of a metallic core surrounded by a rocky mantle, which in turn is surrounded by a thin, rocky crust. Much of the crust is covered by an ocean of liquid, salty water. Surrounding it all is the Earth's atmosphere. Of the rocky and metallic material at or below the surface, only the crust, and in a few locales samples of mantle, are available for direct laboratory study. The composition of most of the Earth's interior is inferred by indirect means, primarily from the study of meteorites.

OVERVIEW

About 4.5 to 4.6 billion years ago, the solar system was formed from a cloud of gas and dust called the solar nebula. The cloud contracted gravitationally, with most of it forming the Sun. The planets and other objects that today orbit the Sun formed by condensation and accretion in an equatorial disk that developed around the early proto-Sun. Small, solid grains condensed from the gas as it cooled. As the grains in the equatorial disk orbited the proto-Sun, they collided and stuck together, accreting into small bodies called planetesimals. As the planetesimals collided and grew into protoplanets, their gravitational fields increased, so they swept up more material in the equatorial disk. The innermost planets—Mercury, Venus, Earth, and Mars—were formed mainly from dense metals and rocks, while the outer planets—Jupiter, Saturn, Uranus, and Neptune—were formed mostly of gases and volatile ices. During or shortly after Earth's accretion, differentiation occurred; the denser metals, such as iron and nickel, sank to the core of the early Earth, while the less dense rocky material rose to the outer portions of the planet.

Samples of the Earth's crust are readily available. Geologic processes have brought samples from the upper part of the mantle to the Earth's surface in certain locales. Most of the Earth's interior is inaccessible to direct study, but meteorites offer clues to its composition. Most meteorites are remnants of the earliest period of planetary formation. They are classified into three main groups based on composition: stony meteorites, stony-iron meteorites, and iron meteorites. Stony meteorites comprise the most abundant group and are composed of silica-associated, or lithophile, elements such as those found in the Earth's crustal materials. Stony-iron meteorites are composed of roughly equal parts of rock (typically the mineral olivine) suspended in a matrix of iron. Iron meteorites are composed of iron (about 80 to 90 percent) along with siderophile elements such as nickel.

Iron meteorites are particularly suggestive to scientists when they attempt to model the composition of the Earth's core. The average density of the entire Earth is about 5.5 grams per cubic centimeter, while the average density of crustal rocks is only about 2.7 grams per cubic centimeter for continental crust and 3.0 grams per cubic centimeter for oceanic crust. This simple comparison indicates that the core must be substantially denser than the average for the entire Earth, and the only reasonably abundant element with about the right density is iron.

The core has two parts: a solid inner core, with a radius of 1,300 kilometers and a density of about 12 to 13 grams per cubic centimeter, and a molten outer core, 2,200 kilometers thick, with a density of about 10 grams per cubic centimeter. The inner core is mostly iron and nickel under high pressure (to make it solid), while the molten outer core probably contains, besides iron and nickel, lighter elements such as sulfur, silicon, oxygen, carbon, and hydrogen. As a whole, the core comprises about one-sixth of the Earth's volume and about one-third of the Earth's mass.

Almost all the remaining two-thirds of the Earth's mass is contained in the mantle, making the mass of the crust, oceans, and atmosphere insignificant in comparison. The mantle is rich in dense, ultramafic rocks such as peridotite, composed mostly of the minerals olivine and pyroxene.

Shortly after (or perhaps during) the initial condensation of grains and their accretion into planetesimals and protoplanets, the Earth's thermal history began through the process of radioactive decay. During this early thermal

Chemical Composition of Earth's Crust

Element	*Weight (%)*	*Volume (%)*
Oxygen (O)	46.59	94.24
Silicon (Si)	27.72	0.51
Aluminum (Al)	8.13	0.44
Iron (Fe)	5.01	0.37
Calcium (Ca)	3.63	1.04
Sodium (Na)	2.85	1.21
Potassium (K)	2.60	1.88
Magnesium (Mg)	2.09	0.28
Titanium (Ti)	0.62	0.03
Hydrogen (H)	0.14	—

period, radioactive nuclides (atoms of specific isotopes) decayed, producing substantial heat that led to at least partial melting. Much of the heating is attributable to the decay of potassium 40 and short half-life elements such as aluminum 26. After as little as perhaps 100,000 years, the planet separated into the iron-nickel core and magnesium-iron-silicate lower mantle. Over a longer timescale (probably more than ten million years but no more than a few hundred million), the high-volatility compounds (such as lead, mercury, thallium, bismuth, water in hydrated silicates, carbon-based organic compounds, and the noble gases) all migrated to the surface, where the material was outgassed or melted into magmas in a continuous period of crustal reprocessing that lasted for several hundred million years.

Separated into three main layers—the core, mantle, and crust—the Earth is an active body, its internal heat far from exhausted. The complexity of the chemical composition increases with each successive outward layer. This generalized model gives a framework for examining the relationships of Earth materials.

Earth's wide range of pressure and temperature regimes helps explain why several thousand distinct minerals and numerous rock types composed of different combinations of minerals have been recognized in samples of the crust and upper mantle. Sampling a variety of crustal rocks leads to a determination of elemental abundance in the crust. By mass, approximately one-half of the crust is oxygen and approximately one-fourth is silicon. These two elements, plus aluminum, iron, calcium, sodium, potassium, and magnesium, make up more than 99 percent of the Earth's crust. Silicon and oxygen combine to form the silicon-oxygen tetrahedron, consisting of a single silicon atom surrounded by four oxygen atoms evenly spaced around it three-dimensionally at the corners of a tetrahedron. This silicon-oxygen tetrahedron joined to additional tetrahedra and/or atoms of other elements forms the class of minerals called silicates, by far and away the most common minerals in the crust.

As ultramafic magmas cool, successive minerals crystallize and settle out via reaction series. As the temperature drops in the melt zone, a discontinuous series (a set of discrete reactions) can occur. Magnetite, an oxide of iron and titanium, is the first to settle out, at about 1,400° Celsius (1,700 kelvins). Olivine, a silicate mineral with a crystal lattice structure of individual silicon-oxygen tetrahedra joined together by other ions (commonly iron and magnesium) and a density between 3.2 and 4.4 grams per cubic centimeter, is the next to crystallize out of the melt. Then comes pyroxene, a silicate mineral with its silicon-oxygen tetrahedra connected in long single chains and a density of 3.2 to 3.6 grams per cubic centimeter. As temperatures in the magma drop to near 1,000° Celsius (1,300 kelvins), the next to crystallize is amphibole, a silicate mineral with its silicon-oxygen tetrahedra joined in long double chains and a still lower density of 2.9 to 3.2 grams per cubic centimeter. As the cooling progresses, the lattice structures increase in complexity with biotite mica, with its silicon-oxygen tetrahedra joined in planar sheets. Paralleling this discontinuous series of reactions is the continuous reaction series of plagioclase feldspar. It has a full three-dimensional lattice of silicon-oxygen tetrahedra, and it varies continuously from being calcium-rich at high temperatures of crystallization to sodium-rich at lower temperatures. Finally, at still lower temperatures down to about 1,000 kelvins (700° Celsius), come potassium feldspar, muscovite mica, and quartz.

With this information, one can start to hypothesize about how the crust and its ocean basins and continents evolved. The oldest Earth materials yet identified are zircon crystals, possibly dating back 4.4 billion years, found in the Jack Hills area of Australia, while the oldest known continental rocks—the Acasta gneiss from the Northwest Territories of Canada—are about 4 billion years old, and they were metamorphosed from earlier igneous rocks. This means that within a few hundred million years after the initial formation of the Earth through condensation, accretion, and differentiation, the first crustal rocks of the Archean eon formed. They probably were composed of olivine, pyroxene, and anorthite (calcium-rich plagioclase feldspar), which crystallized out of basaltic magmas that rose to the surface and cooled and hardened. The early crust, which may have been similar to the anorthosite that makes up much of the ancient highlands on Earth's moon, formed a sheet that was fractured into pieces and subjected to heating through radioactive decay. Differentiation led to the formation of thicker granitic regions surrounded by the thinner basaltic crust. This was the beginning stage in the development of today's crust, which consists of two main types: the denser, thinner, mafic or basaltic oceanic crust and the less dense, thicker, felsic or granitic continental crust. The onset of plate tectonics moved the early continental fragments, causing them to collide and weld themselves together into continental shields in episodes of mountain-building, called orogenies.

The Earth's original inventory of gases appears to have been lost very early in its history, to be replaced with a secondary atmosphere through volcanic outgassing and perhaps impacts of volatile-rich cometary nuclei and carbonaceous chrondrite meteorites. Extensive volcanic activity and high surface temperatures gradually diminished until the hydrosphere (water cycle) was established and oceans appeared.

Life on Earth existed at least 3.5 billion years ago, as evidenced by microfossils similar to modern cyanobacteria (blue-green algae). With the oceans growing in volume and salinity and the development of oxygen-releasing life-forms, Earth's geochemistry became more complex. By the beginning of the Paleozoic era, about 540 million years ago, the oxygen content of the atmosphere had reached 1 percent of its present level. Life-forms significantly shaped the Earth's chemical composition. Multicelled

Primary Rocks and Minerals in Earth's Crust

Rocks	*% Volume of Crust*	*Minerals*	*% Volume of Crust*
Sedimentary		Quartz	12
Sands	1.7	Alkali feldspar	12
Clays and shales	4.2	Plagioclase	39
Carbonates (including salt-bearing deposits)	2.0	Micas	5
		Amphiboles	5
		Pyroxenes	11
Igneous		Olivines	3
Granites	10.4	Clay minerals (and chlorites)	4.6
Granodiorites, diorites	11.2	Calcite (and aragonite)	1.5
Syenites	0.4	Dolomite	0.5
Basalts, gabbros, amphibolites, eclogites	42.5	Magnetite (and titanomagnetite)	1.5
Dunites, peridotites	0.2	Others (garnets, kyanite, andalusite, sillimanite, apatite, etc.)	4.9
Metamorphic			
Gneisses	21.4		
Schists	5.1		
Marbles	0.9	**Totals**	
		Quartz and feldspars	63
Totals		Pyroxene and olivine	14
Sedimentary	7.9	Hydrated silicates	14.6
Igneous	64.7	Carbonates	2.0
Metamorphic	27.4	Others	6.4

Source: Michael H. Carr et al., *The Geology of the Terrestrial Planets*, NASA SP-469, 1984. Data are from A. B. Ronov and A. A. Yaroshevsky, "Chemical Composition of the Earth's Crust," American Geophysical Union Monograph 13.

Typical Composition of Rocks That Compose Much of the Earth's Mantle or Crust

Oxide	*Unmelted Peridotite in the Mantle*	*Basalt Formed at Oceanic Ridges or Rises*	*Andesite Formed at Subduction Zones*	*Granite Rock Along Continental Subduction Zones*	*Continental Rift Basalt*	*Shale*	*Sandstone Near the Source*	*Sandstone Far from the Source*	*Limestone*
SiO_2 (silicon oxide)	45.0	49.0	59.0	65.0	50.0	58.0	67.0	95.0	5.0
TiO_2 (titanium oxide)	0.4	1.8	0.7	0.6	3.0	0.7	0.6	0.2	0.1
Al_2O_3 (aluminum oxide)	8.7	15.0	17.0	16.0	14.0	16.0	14.0	1.0	0.8
Fe_2O_3 (ferric iron oxide)	1.4	2.4	3.0	1.3	2.0	4.0	1.5	0.4	0.2
FeO (ferrous iron oxide)	7.5	8.0	3.3	3.0	11.0	2.5	3.5	0.2	0.3
MnO (manganese oxide)	0.15	0.15	0.13	0.1	0.2	0.1	0.1	—	0.05
MgO (magnesium oxide)	28.0	8.0	3.5	2.0	6.0	2.5	2.0	0.1	8.0
CaO (calcium oxide)	7.0	11.0	6.4	4.0	9.0	3.0	2.5	1.5	43.0
Na_2O (sodium oxide)	0.8	2.6	3.7	3.5	2.8	1.0	2.9	0.1	0.05
K_2O (potassium oxide)	0.04	0.2	1.9	2.3	1.0	3.5	2.0	0.2	0.3
Volatiles (water or carbon dioxide)	1.0	1.0	1.0	2.0	1.0	8.0	2.0	1.0	42.0

Note: Compositions are given as weight percentages of the element oxide in the entire rock.

animals in the oceans scrubbed carbon dioxide from the atmosphere and locked it up in the carbonate rocks, forming biochemically precipitated limestones. By the latter part of Paleozoic era, about 300 million years ago, coal formed as a result of the first land forests being periodically inundated by ocean transgressions.

METHODS OF STUDY

Perhaps no other Earth science is as speculative as that of early Earth history and the geochemical evolution of the Earth. Some of the major challenges confronting Earth scientists are questions about how the Earth's crust formed and when plate tectonic movement began. It is generally accepted by most Earth scientists that heat flow was substantially greater and hence crustal formation occurred more rapidly in Archean times. Despite the problems of extrapolating back to a time when the first solid rocks were forming, the established models are based on some solid lines of evidence.

In 1873, American geologist James D. Dana made one of the initial advances in the study of the Earth's internal chemical composition when he suggested that analogies could be drawn from the study of meteorites. Geochemists studying meteorites today have derived radiometric dates of 4.4 to 4.6 billion years for many of them—corresponding to the initial epoch of condensation and accretion in the solar nebula. Because meteorite types approximate the elemental distribution in the Earth, they are valuable samples of what the Earth formed from.

Geophysicists use seismic waves from earthquakes to study the structure of the Earth's interior. Variations in speed as the waves pass through the Earth, and reflection and refraction of them at internal boundaries, have revealed a differentiated Earth with a very dense metallic core, a less dense rocky mantle, and an even less dense rocky crust "floating" on top. The well-established theory of plate tectonics holds that the crust and upper mantle together form rigid lithospheric plates that are moving, driven by slow convection currents in the mantle.

The drive to study Archean rocks was partly fueled by the United States Apollo missions to the Moon, which returned rocks of comparable age from the lunar surface. Interest in Archean crustal evolution was further aroused by the discovery of Archean lavas called komatiites around greenstone belts (which are agglomerations of Archean basaltic, andesitic, and rhyolitic volcanics, along with their sediments derived by weathering and erosion). Komatiites are ultramafic lavas that formed at temperatures greater than about 1,100° Celsius (1,400 kelvins) and may be fragments of the first crust. Work by field geologists in regions with exposed Archean rocks found successively older granitic rocks—3.8 billion years in western Greenland, 3.9 billion years in Antarctica, and 4.0 billion years in Canada's Northwest Territories. Even older detrital zircons with radiometric ages between 3.8 and 4.4 billion years were discovered in somewhat younger sedimentary rocks in western Australia. The zircon find is significant because it sets an approximate birth date for early continental crust, as zircon is a reasonably common though minor constituent of granitic igneous rocks. The Australian zircons probably formed in early continental igneous rocks and then were eroded, transported, and deposited with other sediments in the sandstones in which they were found.

Geochemists have refined their study of these ancient rocks with more sophisticated methods to determine isotope ratios in them. Instruments common in geochemical laboratories today use X-ray diffraction and gamma-ray spectral analysis to determine which isotopes are present. Isotope ratios in rocks are of particular interest to geochemists because they provide clues as to chemical cycles in nature. The equilibria of these cycles, as indicated by the isotope ratios, offer insights into volcanic, oceanic, biological, and atmospheric cycles and conditions in the past.

CONTEXT

Perhaps no other area of scientific study is as intriguing and controversial as that of the origin and evolution of the Earth. Geochemists and geophysicists have been at the forefront of the quest to understand the Earth's present geology in terms of its past. Before the 1960's, little was known of the Earth's history during early Precambrian times. This lack is significant when one considers that the Precambrian comprises about eight-ninths of the geologic timescale.

It is likely that improved techniques used to analyze rocks and minerals in the laboratory will continue to provide a better understanding of the formation of the Earth's crustal materials and the evolution of moving lithospheric plates. Radiometric dating and isotope analysis will help unravel the relationships between the greenstone belts and granulite-gneiss associations that typify Archean formations on all continents.

Studying features and materials on other solar system bodies will also lead to a better understanding of the early Earth and its evolution. Similarities and differences in Earth's early history are expected to be revealed by future space probes to the Moon, Mars, Venus, Mercury, and asteroids.

David M. Schlom

FURTHER READING

Fyfe, W. S. *Geochemistry*. Oxford, England: Clarendon Press, 1974. Part of the Oxford Chemistry series, this work was written for lower-division college chemistry students. Although in some respects dated, it is nevertheless a brief (about one-hundred-page) and excellent introduction to the science of geochemistry. Of special interest is chapter 9, "Evolution of the Earth." Bibliography, glossary, index.

Gregor, C. Bryan, et al. *Chemical Cycles in the Evolution of the Earth*. New York: John Wiley & Sons, 1988. A systems approach to geochemistry, this book is suitable for the serious college student. Although filled with graphs, tables, and chemical equations, sections are still accessible to the layperson as well. Discussions of mineralogical, oceanic, atmospheric, and other important chemical cycles are extensive and the work is well referenced.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college textbook beyond the introductory level, its approach is based on comparative planetology. Provides much information on condensation and accretion in the solar nebula, as well as the composition of the terrestrial planets generally and Earth in particular.

Kroner, A., G. N. Hanson, and A. M. Goodwin, eds. *Archaean Geochemistry: The Origin and Evolution of the Archaean Continental Crust*. Berlin: Springer, 1984. A collection of reports by the world's leading geochemists studying the geochemistry of the world's oldest rocks. Although many of the articles are technical in nature, the abstracts, introductions, and summaries are accessible to a college-level reader interested in the work of top international scientists.

Levin, Harold L. *The Earth Through Time*. 5th ed. Fort Worth: Saunders College Publishing, 1996. A thorough and readable college text on historical geology. Filled with illustrations, photographs, and figures, this book is also suitable for the layperson. Chapters on planetary beginnings, origin and evolution of the early Earth, and plate tectonics are of special interest. Contains an excellent glossary and index.

Salop, Lazarus J. *Geological Evolution of the Earth During the Precambrian*. New York: Springer-Verlag, 1983. A top Soviet geologist conducts an exhaustive survey of Precambrian geology. Suitable for a college-level reader with a serious interest in the subject. Contains numerous graphs and tables, with extensive references.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. It has a full chapter on the origin, historical development, and composition of the Earth.

Wedepohl, Karl H. *Geochemistry*. New York: Holt, Rinehart and Winston, 1971. An older but still good and accessible brief introduction to geochemistry fundamentals. Contains an excellent chapter on meteorites and cosmic abundances of the elements. Suitable for the nontechnical reader, with index and references. A good starting point for those unfamiliar with mineral formation.

See also: Comets; Earth-Moon Relations; Earth-Sun Relations; Earth System Science; Earth's Age; Earth's Atmosphere; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Structure; Lunar History; Lunar Interior; Planetary Interiors; Planetary Tectonics; Planetology: Venus, Earth, and Mars; Solar System: Element Distribution; Terrestrial Planets.

Earth's Core

Category: Earth

The core is the Earth's densest, hottest region. The thermal energy released by the core's continuous cooling stirs the overlying mantle into slow, convective motions that drive plate tectonics and hence are ultimately responsible for moving continents, building mountains, fueling volcanoes, and producing earthquakes.

OVERVIEW

The Earth's core extends from a depth of about 2,900 kilometers down to the center of the Earth, 6,378 kilometers below the surface at the equator. The outer part of the core is molten, while the central, inner part is solid. Ambient pressures inside the core range from about 1.4 million to 3.6 million atmospheres, temperatures range from about 4,300 to 5,800 kelvins, and densities range from about 10 to 13 grams per cubic centimeter. Being about twice as dense as the rest of the planet, the core contains one-third of the Earth's mass but occupies a mere one-sixth of its volume.

Surrounding the core is the mantle. The boundary between the solid mantle and the underlying molten outer core is the core-mantle boundary (CMB), a surface that demarcates the most fundamental compositional discontinuity in the Earth's interior. Below it, the core is mostly made of iron-nickel alloys. Above it, and all the way to the surface, the mantle and overlying crust are made mostly of silicates (the most abundant group of rock-forming minerals). The core has lower seismic-wave transmission speeds and higher densities than the mantle, a consequence of it having a different chemical composition. The molten outer core probably contains about 80 to 90 percent (by mass) iron-nickel alloys with a 20 to 10 percent mix of sulfur, silicon, oxygen, and maybe even hydrogen. The solid inner core contains less of each of the lighter elements and may be almost entirely iron-nickel alloys that solidified out from the molten outer core. The boundary between the molten outer core and the solid inner core is known as the inner core boundary (ICB); it appears sharp to seismic waves, which easily reflect off it. The entire core must be a good electrical and thermal conductor because of its metallic composition. The mantle, in contrast, is composed mainly of crystalline silicate minerals of magnesium and iron, and therefore is a good electrical and thermal insulator.

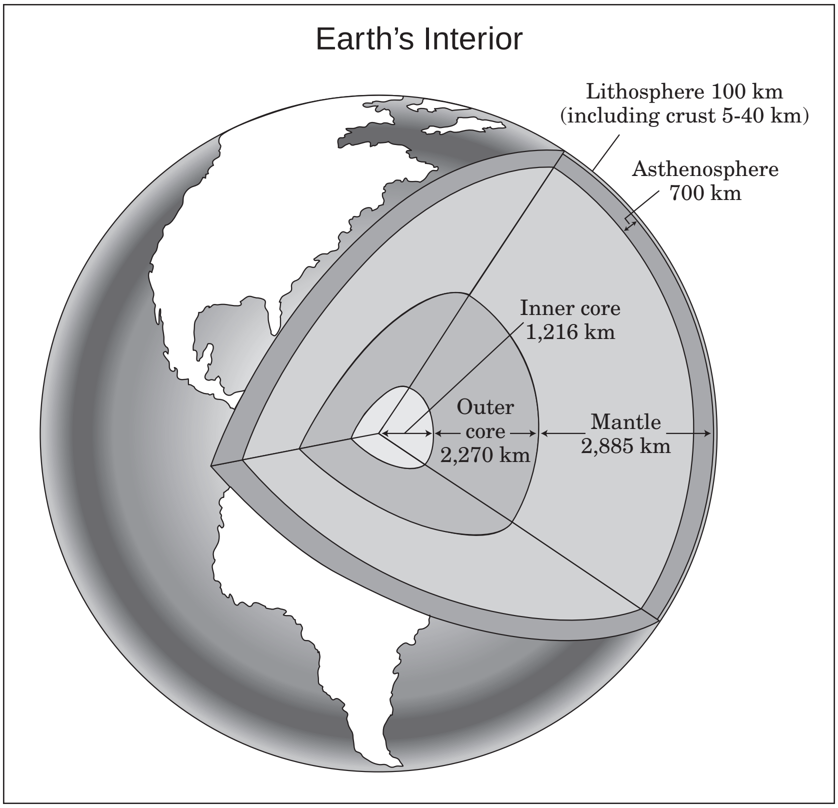
This sharp contrast in physical properties is the end product of the way in which the Earth evolved thermally, gravitationally, and chemically. The Earth formed by the accretion of planetesimals about 4.5 billion years ago along with the rest of the solar system. As the early Earth was slowly heated by radioactivity, the iron in it suddenly melted and sank by gravity toward the center in a cataclysmic "iron catastrophe," forming the core; silicate minerals were left behind to form the mantle and crust. Calculations show that iron sinking to the core must have released great amounts of energy that would have heated and at least partially melted the entire Earth. Cooling of the outer parts proceeded rapidly by convection, but the silicate mantle created a thermal barrier for the iron-rich core, which, not being able to cool down as readily, remained molten. The inner core then began to form at the center, where the pressure was greatest and solidification was (barely) possible.

The most tangible consequence of the existence of an electrically conducting, fluid outer core is the presence of the Earth's magnetic field, which has existed for at least 3.5 billion years with a strength not very different from what it has today. There can be no permanently magnetized substances deep inside the Earth. Magnetic materials lose their magnetism as their temperature increases above the Curie temperature (around 800 kelvins for most magnetic substances), and the interior below a depth of about 30 kilometers is at temperatures well above the Curie point. The process that generates and maintains the geomagnetic field is attributed to a self-exciting dynamo mechanism—that is, an electromagnetic induction process that transforms the motions of the conducting fluid into electric currents, which in turn induce a magnetic field that strengthens the existing field. In order for the system to get started, at least a small magnetic field (perhaps a weak primordial interplanetary field) needs to be present to initiate the generation of electric currents that induce a stronger magnetic field. The increased magnetic field in turn induces stronger currents, which further strengthen the field, and so on. As the magnetic field increases beyond a certain point, it begins to affect the fluid flow through the mechanical Lorentz force, which is induced in a conductor as it moves across a magnetic field. The stronger the magnetic field, the stronger the Lorentz force becomes and the more it tends to modify the motion of the fluid so as to oppose the growth of the magnetic field. The result is a self-regulating mechanism which, over time, attains a steady state.

A source of energy is required to maintain the fluid flow in the molten outer core that produces the electrical currents that induce the magnetic field. One possibility is that the necessary energy to maintain the flow is provided by the growth of the solid inner core as it is fed by the crystallization of iron from the molten outer core. This process could provide enough gravitational energy to keep the core hot throughout, thus driving convection in the molten outer core both thermally and compositionally.

A most extraordinary feature of the core-generated magnetic field is that over the past few hundred million years at least, it has reversed its polarity with irregular frequency, the intervals between reversals varying from less than 100,000 to more than 10 million years. (For convenience, the present orientation of the field is considered normal.) Some igneous rocks and some sedimentary rocks, if they contain iron-bearing mineral grains, can acquire and preserve the magnetism that exists at the time and place where they form. Therefore, rocks formed throughout geologic time have recorded the alternating pattern of normal and reversed Earth magnetism. This sequence of magnetic reversals contains clues to the core's nature.

Geophysicists are eager to determine whether the outer core is vigorously convecting as a consequence of the inner core's growth. If that were the case, convection would be delivering a large flow of heat to the mantle, whose low thermal conductivity would create a barrier to the outward flow of the heat. As a result, the local temperature gradient at the base of the mantle would probably be very high, so that a layer there would be gravitationally unstable. From this layer, thermal inhomogeneities would rise through the mantle in the form of plumes of buoyant, hot, lower-mantle material. Several such plumes might reach the upper mantle or set the entire mantle into convection. These convection currents would be responsible for the motion of the tectonic plates on the Earth's surface and, consequently, for



the drifting of continents, the uplifting of mountain ranges, the formation of oceanic basins, and the occurrence of volcanic eruptions and earthquakes. These observable effects of plate tectonics thus are linked to the internal cooling of the Earth and the growth of the inner core.

METHODS OF STUDY

Knowledge of the structure, physical properties, and composition of the core is entirely based on indirect evidence gathered mostly from analyzing seismic waves, studying the Earth's gravitational and magnetic fields, and conducting laboratory experiments on the behavior of rocks and minerals at high pressures and temperatures. The first evidence for the existence of the core was presented in a paper suggestively titled "The Constitution of the Interior of the Earth, as Revealed by Earthquakes," published in 1906 by Richard D. Oldham, of the geological survey of India. Thirty years later, Inge Lehmann, from the Copenhagen seismological observatory, presented seismic evidence for the existence of the inner core. In the past few decades, with the advent of high-speed computers and technological advances in seismometry, seismologists have developed increasingly sensitive instrumentation to record seismic waves worldwide and sophisticated mathematical techniques that allow them to construct models of the core that explain the observed data.

Seismic waves provide the most important data about the core. Earthquakes and explosions generate elastic body waves that propagate throughout the Earth. These seismic waves penetrate deep into the Earth and, after being reflected or transmitted, travel back to the surface to be recorded at seismic stations around the world. The most direct information that seismic or elastic waves carry is their travel time. Knowing the time it takes for elastic waves to traverse some region of the Earth's interior allows the calculation of their speed of propagation in that region. The speed of seismic waves strongly depends on the density and rigidity of the material through which they propagate, so estimates of the mechanical properties of the Earth's interior can be derived from seismic travel time analyses. Seismic waves that propagate through the interior of the Earth are of two types: compressional waves (also called primary or P waves) and transverse shear waves (also called secondary or S waves). Compressional waves produce volume changes in the medium they travel through; shear waves produce shape distortion without volume change. If the medium has some rigidity (that is, if it is solid), both P and S waves can propagate through it. If the medium has no rigidity (that is, if it is liquid or gaseous), it offers no resistance to a change in shape, so S waves cannot propagate through it, although P waves can.

S waves are not transmitted through the outer core. Therefore, the outer core material has no rigidity, but behaves as a fluid would. Similar observations suggest that the inner core is solid. However, the actual rigidity of the inner core is very difficult to estimate since shear waves inside the inner core are isolated from the mantle by the outer core and can only travel through it as P waves converted from S waves at the inner core boundary. Nevertheless, when the whole Earth is set into vibration by a very large earthquake, the average rigidity of the inner core can be estimated by comparing the observed frequencies of oscillation with those theoretically computed for models of the Earth that include a solid inner core. Model studies have indicated that the inner core is indeed solid, because a totally liquid core model does not satisfy the observations.

The average speed of P waves through the entire Earth is about 10 kilometers per second, whereas the P-wave speed in the rocks accessible to measurement at the Earth's surface is about 4 to 5 kilometers per second. The speed of S waves is a little more than half that of P waves in solids and is zero in perfect fluids. The P-wave speed drops abruptly across the core-mantle boundary, from 13.7 kilometers per second at the base of the mantle to 8.06 kilometers per second at the top of the core. From this point down, the speed steadily increases to 10.35 kilometers per second at the inner core boundary, where it jumps discontinuously to 11.03 kilometers per second at the top of the inner core. From there to the center of the Earth, the P-wave speed increases slowly to reach 11.3 kilometers per second. The S-wave speed increases from zero at the inner core boundary to around 3.6 kilometers per second at the Earth's center. The core's density abruptly increases from 5,500 kilograms per cubic meter at the base of the mantle to nearly 10,000 kilograms per cubic meter at the top of the core. From there, the density increases slowly to nearly 13,100 kilograms per cubic meter at the Earth's center. For comparison, the density of mercury at room temperature and ambient pressure is 13,600 kilograms per cubic meter.

The theory that the core is mostly iron is consistent with iron being cosmically more abundant than other heavy elements and with the high electrical conductivity the core needs in order to generate the Earth's magnetic field. The fluidity of the outer core has been demonstrated by measurements not only of seismic wave transmission but also of the oscillation period of gravitational waves in the core excited by the lunisolar tides. The existence of a sustained magnetic field is also consistent with a fluid outer core.

Seismic data can probe the inner core only partially from the Earth's surface, unless the source of the seismic waves and the receivers are located antipodally to each other. Such an arrangement allows scientists to measure seismic waves that have penetrated the center of the Earth.

New views of the Earth's interior are produced, sometimes unexpectedly, by the analyses of data collected by satellite missions. Data from orbiting satellites that measure tiny variations of the Earth's gravitational field, combined with computer-aided seismic tomography of the Earth's interior, have revealed large-density anomalies at the base of the mantle and a large relief of more than 2 kilometers along the core-mantle boundary.

Seismic tomography uses earthquake-generated waves that penetrate the interior in a multitude of directions to map its three-dimensional structure, just as computerized medical tomography uses multiple X-ray images to create a three-dimensional view of internal organs of the body. Essential to the success of these studies is installation of dense networks of seismic sensors all over the surface of the Earth; this installation, however, would be very expensive.

CONTEXT

Any study of the Earth provides insight into the nature and future of the planet and, consequently, the future of humankind. A more complete understanding of how the Earth's core works could result in predictions of the geomagnetic field's activity for years to come, including perhaps an upcoming magnetic reversal. According to the best estimates, a reversal does not occur instantaneously, but takes at least a few thousand years. That means that during a reversal, there is a time interval of very weak or even zero field intensity. Under such conditions, the magnetic shielding that prevents highly energetic charged particles of the solar wind from reaching the Earth's surface would disappear, leaving the surface directly exposed to intense particle radiation that could directly be potentially lethal, or that could at least lead to increased cancers and genetic mutations.

The inner core has not yet been thoroughly explored with seismic waves. One reason is that it is the most remote region of the Earth and therefore the most difficult to study; another is that it is hidden beneath the "seismic noise" created by the crust, mantle, and outer core. The inner core, however, holds the key to the understanding of the Earth's early history and its subsequent development as a planet. It would be possible in principle to investigate the inner core in more detail by deploying an array of highly sensitive seismic sensors antipodal to a seismically active region. However, despite the wealth of unique data that could be obtained from such an experiment, it would be a very expensive endeavor.

J. A. Rial

FURTHER READING

Bolt, Bruce A. *Inside the Earth: Evidence from Earthquakes*. San Francisco: W. H. Freeman, 1982. An elementary treatment of what is known about the Earth's interior, mostly through the study of seismic waves, the author's major field of research. The book contains abundant diagrams that illustrate important results of the investigation of the core and mantle. For readers with some knowledge of mathematics, the book includes brief derivations of important formulas, separated by "boxes" from the main text. Includes anecdotal descriptions of great scientific discoveries along with personal views of the history and development of seismology.

Fowler, C. M. R. *The Solid Earth: An Introduction to Global Geophysics*. 2d ed. New York: Cambridge University Press, 2004. An updated version of a widely used textbook for introductory geophysics courses. Designed for students with some knowledge of physics and calculus.

Jacobs, J. A. *The Earth's Core*. 2d ed. New York: Academic Press, 1987. A highly technical text, but perhaps the best reference for a detailed description of the most accepted core models. The tables—which give the numerical values of the density, temperature, rigidity, and wave velocity distributions within the Earth—are of interest to anyone wanting a quantitative description of the core. A long list of research articles is included.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. There is a very good chapter on the Earth's interior.

Van der Pluijm, Ben, and Stephen Marshak. *Earth's Structure*. 2d ed. New York: W. W. Norton, 2003. An introductory text on structural geology and tectonics. Designed for undergraduate students.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996. Covers geophysics research and theories since 1960. Includes information about Pangaea, the supercontinent cycle, and the reversals of the Earth's magnetic field. For general audiences.

See also: Earth-Moon Relations; Earth-Sun Relations; Earth's Age; Earth's Composition; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Structure; Planetary Interiors; Planetology: Venus, Earth, and Mars; Terrestrial Planets.

Earth's Core-Mantle Boundary

Category: Earth

The core-mantle boundary is a pronounced discontinuity separating the outer core from the mantle of the Earth. It is a chemical and mineralogical as well as a thermal and mechanical boundary. The topography of the core-mantle boundary is believed to be controlled by the dynamic processes in the mantle and the outer core.

OVERVIEW

The core-mantle boundary (CMB) is a prominent discontinuity within the Earth. It is located at a radius of about 3,500 kilometers and a depth below the surface of about 2,900 kilometers. The mantle above the boundary is largely solid, of relatively low temperature, and primarily composed of silicate minerals rich in magnesium and iron. The outer core below the boundary is liquid, of higher temperature, and composed of dense materials such as iron-nickel oxides and iron-nickel sulfide alloys. This boundary separates two dynamic systems: one operating in the mantle as hot spots and convection cells, the other in the outer core consisting of convection currents and eddies of the core fluid. The motions of the core fluid appear to be responsible for the Earth's magnetic field.

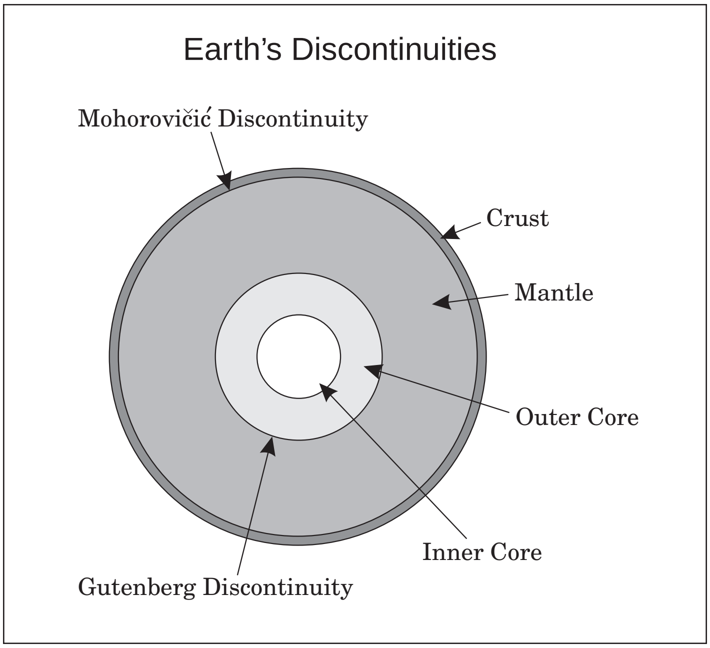
The lowermost part of the mantle, labeled the D″ (pronounced "dee double prime") layer, is called the core-mantle transition zone. It is approximately 200-300 kilometers thick and is located just above the CMB. Seismic waves from earthquakes and explosives detonated at or near the surface show significant speed variations within the D″ layer over lateral or horizontal distances of 1,000 kilometers or more. Longitudinal (or compressional) P waves travel faster in the portions of this layer that are located below North America, China, the eastern part of the Indian Ocean, and off the Pacific coast of Chile; P waves travel more slowly in the D″ layer below the southern part of Africa, the New Hebrides Islands, the South Pacific Ocean, and the Argentine Basin. Similar variations in speed have also been observed for transverse (or shear) S waves, which travel faster in this layer under the American continents, Asia, the northern Indian and Pacific oceans, and Antarctica, and more slowly under the Central and South Pacific Ocean, the Atlantic Ocean, most of Africa, and the southern part of the Indian Ocean. These speed variations in the D″ layer appear to continue upward in the mantle and may be related to the thermally induced convection currents and hot spots in the mantle.

Improvements in instrumentation have enabled scientists to simulate in the laboratory the physical and chemical conditions of the lower mantle and the outermost core. The lower mantle is thought to be composed primarily of magnesium-iron silicates with the compact perovskite crystal structure. Some aluminum-calcium silicates and magnesium-iron oxides may also be present, but their relative abundance is not known. Laboratory measurement of the melting point of perovskite lead to estimates of the temperature of the D″ layer varying from 3,300 to 4,300 kelvins. Similar studies of outer core materials, which are primarily iron-nickel sulfides and iron-nickel oxides, indicate that the temperature of the outermost core is at least 4,300 kelvins. Thus, the temperature increases by about 1,000 kelvins through the D″ layer, resulting in partial melting of some minerals, thereby making the zone soft with anomalous characteristics.

Seismologists studying the core-mantle boundary (CMB) by means of reflected waves from the core have long been frustrated by the strong scatter of the reflected amplitudes. A major part of this scatter is believed to be the result of vertical undulations of the core-mantle boundary. The lateral extent of these undulations is of the order of thousands of kilometers. The elevation of the boundary may change by as much as 5 to 8 kilometers above or below its normal depth. Topographic highs of the CMB have been observed beneath the Indian Ocean, the Pacific Ocean, and the Atlantic Ocean (particularly in the North Atlantic). The CMB is depressed below the Tonga-Karmadec islands, the China-Japan region, Central Africa, and off the west coast of South America. Because most of these areas are associated with the subduction of oceanic plates, the structure of the CMB is thought to be caused by the dynamic processes in the mantle, which may be related to the convection processes in the outer core. Subduction of a lithospheric plate is associated with downwelling convective flow in the mantle. When the convective flow reaches the core boundary, it depresses the CMB into the hot, liquid core. Core fluids may partially invade the "topographic low" of the CMB, altering the chemical composition of the D″ layer. Similarly, beneath an upwelling zone of mantle flow, liquid core material may be "sucked" up into the mantle, creating a topographic high of the CMB. At the topographic "lows" of the CMB, mantle material is subjected to the higher temperatures of the outer core and may melt; it then recrystallizes at the topographic "highs" of the CMB. Thus, the overall effect is to smooth the CMB, which is continually disturbed by the convective circulations in the mantle. With heat dissipation in the mantle, the outer core slowly cools and the core materials crystallize and underplate the mantle. Therefore the outer core slowly shrinks with time, and the CMB gets deeper.

The D″ core-mantle transition zone and core-mantle boundary are not only a compositional boundary but also a thermal boundary, where the temperature increases by at least 1000 kelvins. Thermal coupling between the mantle and the outer core, however, may change laterally, resulting in a variable heat flow across the boundary. Although no consensus has been reached among scientists, it is possible that the mantle dynamics are at least partially responsible for controlling the heat flow.

The Earth behaves like a large magnet. The magnetic field of the Earth—the geomagnetic field—undergoes changes known as secular variations. The origin of the geomagnetic field appears to be related to motions of the outer core fluid. Studies suggest that the deep mantle and the outer core play a significant role in shaping the secular variations. Upwellings in the outer core material are associated with the hotter and seismically slower regions of the D″ zone; downwellings are associated with the colder and seismically faster regions in D″. Cold regions in the mantle transmit greater amounts of heat from the outer core, thereby setting up



mantle circulations. The topography of the CMB is controlled by the circulations in the mantle as well as in the outer core. The topographic relief of the CMB may also set up a lateral temperature gradient which may be responsible for the secular variation of the magnetic field.

METHODS OF STUDY

Various subdisciplines of geophysics contribute to investigating the nature and the structure of the core-mantle boundary (CMB) and the D″ core-mantle transition layer. They include, among others, seismology, geodesy, geodynamics, high-temperature and high-pressure mineral physics, geothermometry, and geomagnetism. Seismology has been the most important among all these subdisciplines and has contributed most of the information about the Earth's interior.

Seismology deals with earthquakes and the propagation of earthquake waves through the Earth. Whenever an earthquake occurs, different types of seismic waves are generated. Surface waves travel along the Earth's surface, and longitudinal (or compressional) P waves and transverse (or shear) S waves travel through the interior of the Earth. It is often helpful to visualize the direction of travel of P and S waves as rays originating from an earthquake focus, or hypocenter, and radiating in all directions through the Earth. Because of the increased rigidity and incompressibility of rocks downward, the speeds of these waves increase with depth. As a result, the downgoing seismic rays (except for vertical or near-vertical rays) are curved back toward the surface. The seismographic stations that are farther away from the epicenter record seismic waves that penetrate through the deeper layers in the Earth.

The outer core has no rigidity, since it is liquid. Consequently, P waves slow down abruptly as they cross the CMB, from 13.5 to 8.5 kilometers per second, and they are sharply refracted or bent. S waves do not propagate through liquids, so they stop at the CMB. As a result, there are shadow zones beyond about 11,000 kilometers from the epicenter where neither P nor S direct waves are detected at the ground surface. The presence of liquid outer core was discovered through the existence of the shadow zones and the absence of core-transmitted S waves.

Seismic waves emerging at steep angles from the hypocenter encounter the CMB. Part of the incident energy is reflected back from the boundary, and the rest is refracted through the outer core. P waves can be reflected back as P and as S waves, designated as PcP and PcS waves (or phases) respectively. Similarly, S waves reflected back from the CMB as P and S waves are designated as ScP or ScS waves. These core-reflected phases have been important in the study of the nature, shape, and depth of the CMB. Because S waves cannot travel through liquids, the refracted energy in the outer core propagates in the form of P waves. These refracted P waves are designated as K phases. Thus PKP is a wave that travels from the hypocenter in the mantle as a P wave, propagates as a K (that is, P) in the outer core, and reemerges in the mantle as a P wave. Similarly, SKS and other combinations, such as PKS and

SKP, are often observed in the seismic records. A joint study of the core-reflected phases (for example, PcP) and the core-refracted phases (for example, PKP) is often important in resolving the depths and topography of the CMB. Seismic rays incident at a large angle on the CMB are diffracted. Study of these diffracted waves provides important information on the D″ zone above the CMB. Using the waveform modeling techniques, scientists are determining the thickness and fine structures of the D″ zone.

Another important tool is seismic tomography. It utilizes the same principle used in computed tomography (CT) scan X rays of humans. In a CT scan, the X-ray source and imager are rotated around the body and a large number of X-ray images are recorded. A computer processes these images and forms a three-dimensional image of the internal organs of the subject. The seismological data collected worldwide can similarly be processed to form a three-dimensional image of the Earth's interior. Seismic tomography provides valuable information on the CMB as well as the Earth's mantle.

A large earthquake sets the Earth vibrating like a bell. If the Earth were perfectly spherical with uniform layering, it would produce a pure tone, vibrating at a preferred frequency. Departures from the spherical shape of the Earth, as well as depth-related discontinuities, produce additional tones involving distortions of the Earth. Thus, recordings of these various modes of the Earth's vibrations, known as free oscillations, can furnish information about the shape of the CMB.

Satellite measurements of the Earth's gravity field and the geodetic observations of the geoid can also provide information on the CMB. Theoretical models of the Earth's interior, particularly the mantle, the D″ zone, and the CMB, can be constructed to match observed geoidal undulations and gravity anomalies. It appears that a 2- to 3-kilometer variation in elevation of the CMB can explain 90 percent of the observed large-scale gravity anomalies. Astronomic observations of the Earth's wobble also furnish additional constraints on the shape of the CMB. The Earth has an equatorial bulge caused by its rotation. The Moon pulls at the bulge and attempts to align it along the orbital plane of the Moon, generating a wobble, or a nutational motion, of the Earth's axis. (This motion is similar to the wobble of a spinning top or a gyroscope.) Deformation of the CMB produces certain irregularities in the nutational motion. Studies of these irregularities indicate that the undulations of the CMB are less than 1 kilometer in height.

Major developments in instrumentation have made it possible to simulate in the laboratory the temperature and pressure conditions of the deep mantle. Scientists can now study how the crystal structures of minerals change with increased temperature and pressure. Measurements of the electrical properties of rocks under high pressure, and possible alloying of iron by sulfur and oxygen that may occur in the outer core, are also being studied. These investigations are important for complete understanding of the mineral compositions, structure, temperature, and pressure environment of the Earth's deep interior.

CONTEXT

Scientists from various geophysical subdisciplines have made a concerted effort to investigate the structure and nature of the CMB and the deep interior of the Earth because it is important from several perspectives. The CMB is believed to be associated with deep mantle plumes, the mantle convection currents that drive the lithospheric plates and may be responsible for secular variations of the geomagnetic field. As the most pronounced discontinuity within the Earth, the undulations at the CMB may also cause regional gravity anomalies and can affect the transmission of seismic waves. The transmission effects of seismic waves crossing the CMB provide information about the geometry and the physical and chemical properties of materials at the CMB, as well as in the mantle above and the outer and inner core below. Furthermore, because core-reflected phases travel along vertical or near-vertical paths in the mantle, they are often utilized to study heterogeneity and seismic behavior in the mantle. Knowledge of the nature of the CMB is necessary to determine these mantle characteristics.

A committee on Studies of the Earth's Deep Interior (SEDI), under the joint auspices of the

International Union of Geodesy and Geophysics (IUGG) and the American Geophysical Union (AGU), facilitates international exchange of scientific information about the Earth's interior. In addition, special sessions on the Earth's deep interior are held at most AGU meetings.

D. K. Chowdhury

FURTHER READING

Bolt, Bruce A. *Earthquakes*. New York: W. H. Freeman, 1988. This volume presents information on the Earth's interior obtained from seismological studies. Suitable for high school and college levels.

_______. *Inside the Earth*. San Francisco: W. H. Freeman, 1982. A good introduction to seismology for the nonscientist, this well-illustrated, concise book summarizes the seismological methods and the results.

Fowler, C. M. R. *The Solid Earth: An Introduction to Global Geophysics*. 2d ed. New York: Cambridge University Press, 2004. An updated version of a widely used textbook for introductory geophysics courses. Designed for students with some knowledge of physics and calculus.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. There is a very good chapter on the Earth's interior, that includes sections on the CMB and the D″ layer.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996. Covers geophysics research and theories since 1960. Includes information about Pangaea, the supercontinent cycle, and the reversals of the Earth's magnetic field. For general audiences.

See also: Auroras; Earth-Moon Relations; Earth-Sun Relations; Earth System Science; Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Core; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Rotation; Earth's Shape; Earth's Structure; Eclipses; Greenhouse Effect; Van Allen Radiation Belts.

Earth's Crust

Category: Earth

The crust is the outermost layer of the Earth. The dynamic changes involved in the creation and destruction of crustal rock also fuel volcanoes, cause earthquakes, concentrate mineral deposits, and liberate gases and water that form the atmosphere and ocean.

OVERVIEW

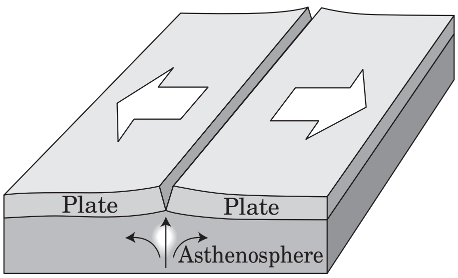
The crust of the Earth is the outermost layer of the Earth. It is distinct from the region of rock lying beneath it, called the mantle, in that the rocks that comprise the crust have different compositions and a lower density. Continental crust is composed mostly of granitic rocks with densities around 2.7 grams per cubic centimeter, and oceanic crust is composed mostly of basaltic rocks with densities around 3.0 grams per cubic centimeter. In contrast, rocks from the upper mantle have densities around 3.3 grams per cubic centimeter, and probably are composed mostly of peridotite.

The rocks of the Earth's crust are quite varied. They can be classified as belonging to one of three broad groups, depending on how they formed: igneous, sedimentary, and metamorphic. These three groups are parts of what is referred to as the rock cycle, which depicts the way in which rocks from each of these groups can provide the raw material to form rocks in any other group.

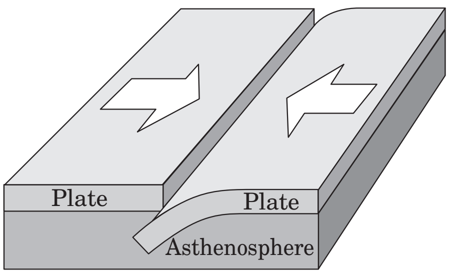
Igneous rocks are formed by cooling and crystallization from molten material called lava (if on the surface) or magma (if below the surface). Igneous rocks that cool and harden on the surface are said to be extrusive, and those that cool and harden below the surface are said to be intrusive. Intrusive igneous rocks cool more slowly and thus usually contain larger mineral crystals (large enough to be seen with the unaided eye). Extrusive igneous rocks cool more rapidly and thus either contain smaller, microscopic mineral crystals or are glassy, containing no crystals at all. Some common igneous rocks are granite and gabbro (intrusive) and rhyolite, basalt, and obsidian (extrusive).

Sedimentary rocks are formed from sediment, the remnants of other rocks that were weathered and eroded when exposed at the surface. Weathering processes may be chemical (such as dissolving in water or acid, or being oxidized by oxygen from the atmosphere) or physical (such as being broken apart when water in cracks freezes and expands, or roots grow into cracks). Sediment can be transported by various agents (such as moving water or blowing wind) and is ultimately deposited in layers. The sediment may then be compacted and cemented, forming sedimentary rock. Some common sedimentary rocks include sandstone, shale, and limestone.

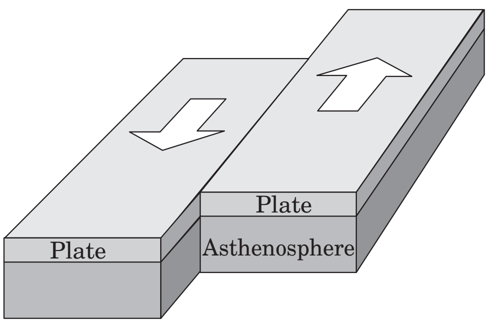
Plate Boundaries



Divergent Boundary



Convergent Boundary



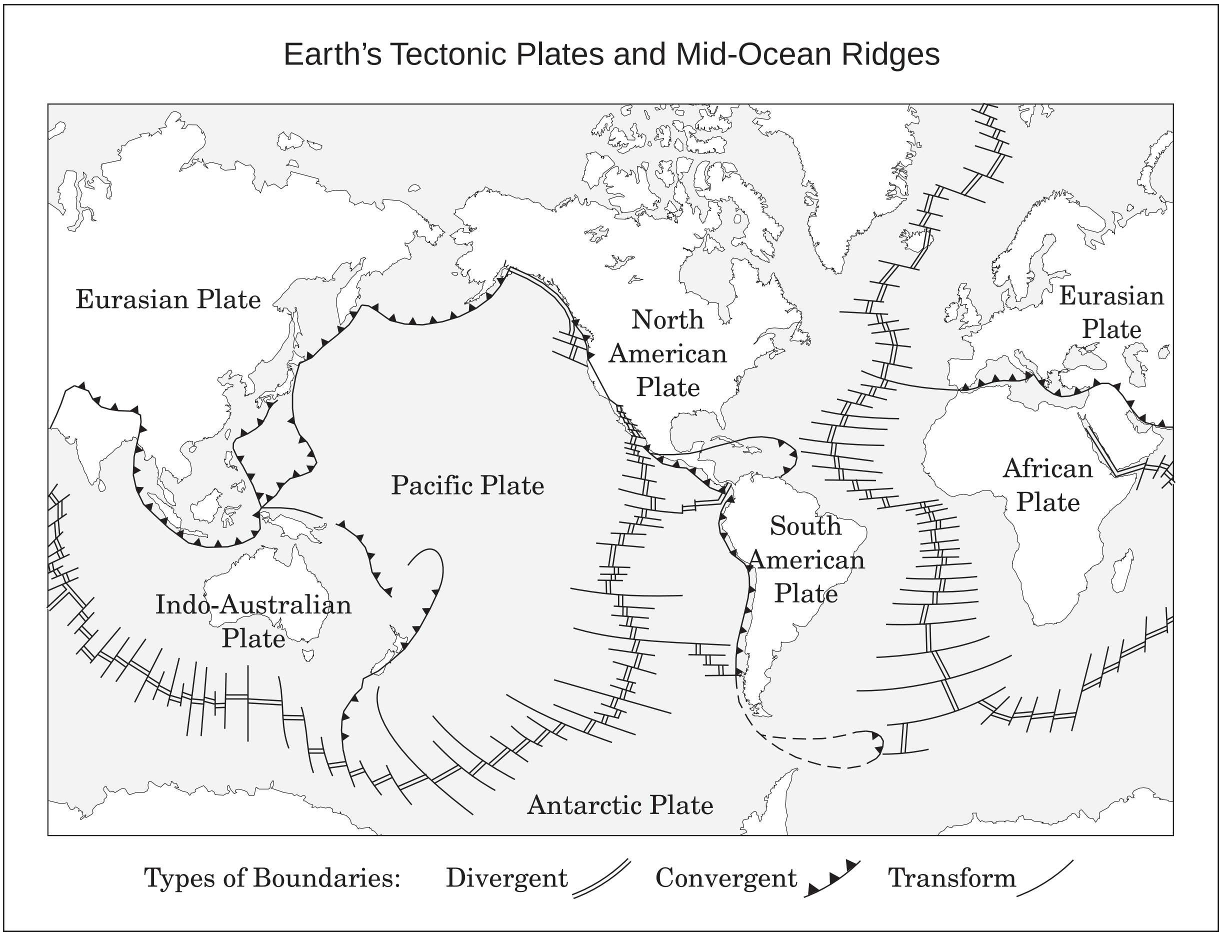
Transform Fault Boundary

Metamorphic rocks are formed from other rocks that have been subjected to pressures and temperatures high enough to change the crystalline structure of the minerals in the rock (but not high enough to melt the rock). Such changes often occur in the deeper parts of the crust, where the temperature and pressure are greater (around 400 to 1,200° Celsius, and 3,000 to 15,000 atmospheres). Some common metamorphic rocks are slate, schist, gneiss, and marble.

The boundary between the crust and the mantle is known as the Mohorovičić discontinuity or simply the Moho. The depth of the Moho, and thus the thickness of the crust, varies widely. The crust is thickest under the continents, averaging about 40 kilometers and reaching a maximum of 70 kilometers beneath young high mountain chains such as the Himalayas. Under the ocean basins, the crustal thickness averages only about 7 kilometers.

The crust and uppermost mantle (down to a depth of about 100 kilometers) constitute the rigid lithosphere, which is divided into a number of separate blocks called plates. These plates are slowly moving, driven by slow convective motions in the soft, plastic part of the mantle called the asthenosphere, which is directly beneath the lithosphere.

The crust is constantly being created, deformed, and destroyed by processes at plate boundaries. There are three types of boundaries between plates: divergent, where plates are pulled apart; convergent, where plates collide; and transform or side-slip, where plates slide horizontally past each other. Divergent plate boundaries, where plates move apart, are



marked by the ocean ridge-rift system where new ocean basin crust is produced by hot upwelling magma. The Mid-Atlantic Ridge (located in the Atlantic Ocean about midway between North and South America to the west and Europe and Africa to the east) and the Red Sea (between Africa and Arabia) are examples of divergent boundaries. Convergent plate boundaries, where plates come together, can be divided into three subtypes, depending on the nature of the plate (oceanic or continental) on each side of the boundary: oceanic-oceanic, oceanic-continental, and continental-continental. At oceanic-oceanic and oceanic-continental convergent boundaries, subduction occurs. Subduction is a process in which an oceanic plate descends back down into the mantle. Subduction zones are marked by deep-sea trenches that are located where oceanic plates bend down into the mantle, there to be heated and remelted, fueling chains of volcanoes parallel to the trench. The Mariana Trench in the western Pacific and the Peru-Chile Trench along the west coast of South America are examples of oceanic-oceanic and oceanic-continental convergent boundaries respectively. At continental-continental convergent boundaries, neither plate subducts, but instead crumple and buckle along the collision zone forming high mountain ranges like the Himalayas between India and the rest of Asia. At transform or side-slip plate boundaries, two plates slide horizontally past each other. The San Andreas fault in California is a classic example of a transform plate boundary.

New ocean basin crust forms at the ocean ridge-rift system, which is located along divergent plate boundaries. As the existing lithospheric plates are pulled apart, probably by very slow lateral flow in the underlying asthenosphere, upwelling magma fills the gap. It cools and hardens, producing the igneous rocks basalt and gabbro, which add new bands of oceanic crust to the plates. The newly formed oceanic crust spreads away from the ocean ridge-rift in a process known as seafloor spreading. The age of oceanic crust increases systematically with distance from the ocean ridge-rift system. The oldest known seafloor crust, about 180 million years old, is found in the western Pacific. Apparently in that sort of time frame or less, oceanic crust reaches a subduction zone and is recycled back into the mantle. In contrast, many continental rocks are several hundred million years old, and the oldest continental rocks found so far, from the Northwest Territories of Canada, are about 4 billion years old, almost as old as the 4.5 billion-year age of the Earth.

Continental crust does not subduct. It remains on or near the surface and is recycled there through the processes and stages of the rock cycle. It is initially created along subduction zones where oceanic crust is consumed. As the lithospheric plate descends back down into the mantle, it is heated. Eventually, differential melting occurs; the minerals with lower melting temperatures melt and rise as molten blobs of magma, either cooling and hardening below the surface as igneous intrusions or erupting at the surface as volcanoes. These igneous intrusions and volcanoes are composed of granite, rhyolite, and similar igneous rocks that are less dense than the basalts and gabbros of oceanic crust. They form continental crust, which, because of its lower density, does not subduct but remains at the surface. When two continental plates collide, they crumple and weld themselves together, forming a high mountain range with roots that extend downward, increasing crustal thickness there. Thus, continents grow with time by two processes occurring along their edges: volcanism near subduction zones and accretion. Continents also can be broken apart when rifts develop in them, like the East African Rift. If the rift continues to grow, eventually it becomes a long narrow arm of the ocean, called a linear sea, like the Red Sea between Africa and Arabia.

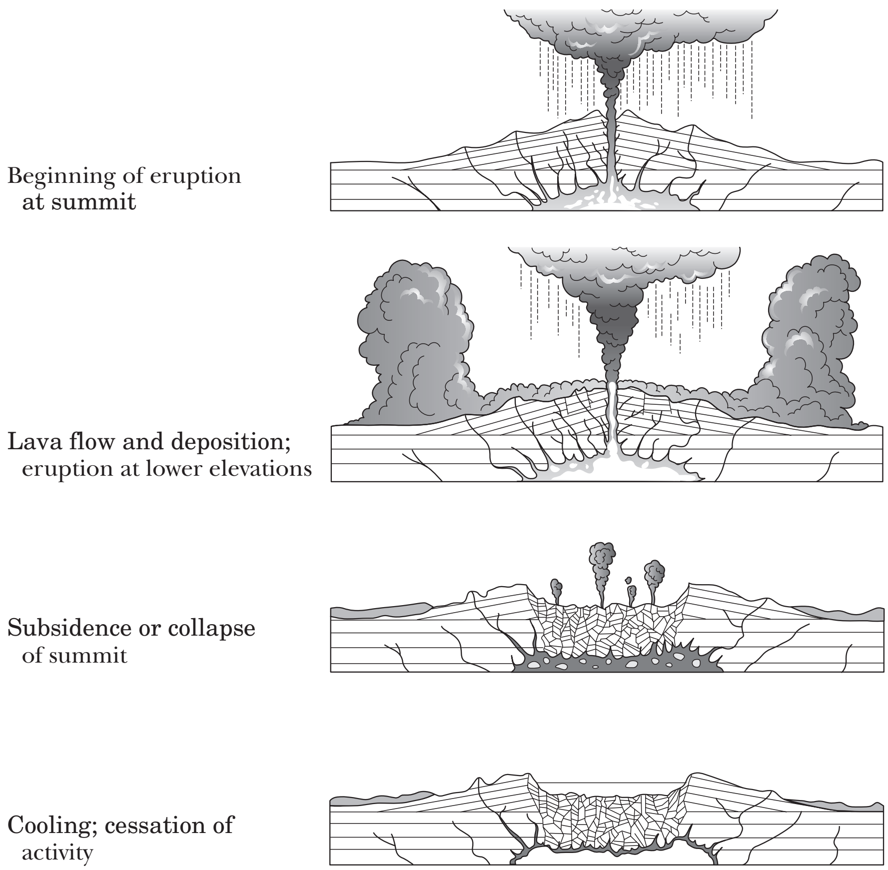
The crust is thickest under young mountain belts, called orogenic belts, piling upward and sinking downward simultaneously to form a thick wedge of rock. In this sense, it is much like a buoyant iceberg, floating with the majority of its mass below the water. The buoyancy of the less dense crustal rocks floating on the more dense mantle rocks is known as isostasy or flotational equilibrium. Just as the iceberg must reach a flotational level by displacing a volume of water equal to its mass, so must the lighter crustal rocks displace a volume of denser mantle rocks to reach their buoyancy level. Thus the thickest crust is found under the highest mountainous regions because they have such deep roots, while the thinnest crust is found under ocean basins.

Toward the center of continental landmasses are large areas of old rocks known as cratons. The ages of rocks found in the cratons range from about 600 million to as much as 4 billion years. The cratons have been free of deformation and mountain-building for at least the last 600 million years. Consequently, their surfaces tend to be relatively flat as a result of surface processes such as weathering and erosion acting on the exposed rocks over geologically long periods of time. Cratons have long and complex histories. Over large areas of them, highly deformed rocks from the deep roots of ancient mountain regions are exposed, and in other regions the ancient mountain roots are covered by more recent layers of sedimentary rocks.

METHODS OF STUDY

Seismic waves created by both earthquakes and artificial explosions are used to probe the interior of the Earth, including its top layer, the crust. One type of seismic wave that travels through the Earth is a longitudinal compressional wave called a P wave (for primary wave); this type of wave is analogous to an acoustic or sound wave, alternately compressing and stretching the material it travels through. The density, rigidity, and compressibility of the material determine the wave speed in it. P waves travel at speeds between about 2 and 6 kilometers per second near the surface, because of the wide range of compositions of surface rocks as well as the presence of open space and fluids

Volcanic Eruption and Caldera Formation



within them. (For comparison, sound waves travel at about 0.3 kilometers per second through the lower atmosphere.) P-wave speed reaches about 6 to 7 kilometers per second in the lower crust just above the Moho. It has been found in the laboratory that metamorphic rocks known as granulites, which can form when basalts and gabbros are subjected to the pressures and temperatures of the lower crust, have P-wave speeds in the proper range. Furthermore, granulites are similar to some rock samples brought to the surface in volcanic pipes that are thought to have originated in the lower crust.

The thickness of continental crust has been determined by the study of seismic waves that are reflected off the Moho, as well as those that refracted by it. The increase in density and the resulting increase in speed of seismic waves when they cross the Moho from crust to mantle cause their wave path to bend or refract. Waves that cross the Moho at what is termed the critical angle will travel along right beneath the Moho and be refracted back to the surface at the same angle. By Snell's law of refraction, the sine of the critical angle equals the ratio between the crust and mantle speeds. The travel time of these refracted seismic waves from their source till they are recorded on a seismograph is determined by the crust and mantle speeds, and the thickness of the crust. Using critically refracted P waves, thicknesses have been estimated for much of the crust. It is possible to check the crustal thickness determined from critically refracted waves by using reflected waves. This has been applied with particular success in continental areas with artificial acoustic wave sources such as explosives and vibrator trucks.

CONTEXT

The Earth's crust is in a state of dynamic evolution, with rock materials being created, deformed, and destroyed. This dynamic evolution of the crust is brought about by the movement of large lithospheric plates composed of the crust and upper mantle, up to 100 kilometers thick. Processes at plate boundaries can produce volcanism and earthquakes.

Volcanic activity over the billions of years of the Earth's existence has provided the water vapor and other gases necessary to form the oceans and atmosphere by the release of gases trapped in lavas that reach the surface. An understanding of volcanoes, including how, why, and where they occur, requires an understanding of the Earth's crust and crustal dynamics. It is especially important in the areas with active volcanoes to be able to assess the hazards they pose.

Plate motion also produces earthquakes where two plates rub against each other. Usually the strongest occur at transform plate boundaries and at subduction zones along convergent boundaries. Eventually, knowledge of how crustal rocks change and respond before impending earthquakes may allow their prediction.

Exploration for important economic minerals is guided by knowledge about the evolution and composition of the crust. The concentration of valuable metal deposits, such as gold and copper, occurs during volcanic activity at ocean ridge sites where new oceanic crustal rocks are being created. Consequently, exploration efforts for such metallic ores can be directed toward identifying ancient ridge site locations. The formation of continental sedimentary rocks in the Gulf of Mexico traps organic materials that will be turned into oil and natural gas. Looking for similar types of sedimentary rocks in ancient crustal environments aids in the search for petroleum and natural gas.

David S. Brumbaugh,
revised by Richard R. Erickson

FURTHER READING

Bally, A. W. *Seismic Expression of Structural Styles*. Tulsa, Okla.: American Association of Petroleum Geologists, 1983. An excellent visual treatment of the structure and layering primarily of the upper crust throughout the world. Sections into the crust of offshore Scotland and northwest Germany show the Moho. Suitable for a broad audience from general readers to scientific specialists.

Bott, M. H. P. *The Interior of the Earth*. New York: Elsevier, 1982. This book was intended for undergraduate and graduate students of geology and geophysics as well as for other scientists interested in the topic. The plate tectonic framework of the outer part of the Earth is strongly emphasized.

Brown, G. C. *The Inaccessible Earth: An Integrated View to Its Structure and Composition*. 2d ed. New York: Chapman and Hall, 1993. A good general introduction geared toward the undergraduate college student. The primary topics are the internal state and composition of the Earth. Included is background material on seismology and three chapters discussing the Earth's crust.

Fowler, C. M. R. *The Solid Earth: An Introduction to Global Geophysics*. 2d ed. New York: Cambridge University Press, 2004. An updated version of a widely used textbook for introductory geophysics courses. Designed for students with some knowledge of physics and calculus.

Knapp, Ralph E. *Geophysics*. Exeter, England: Pergamon Press, 1995. Thorough treatment of the physics of Earth. For the serious student.

Smith, David G., ed. *The Cambridge Encyclopedia of Earth Sciences*. New York: Cambridge University Press, 1982. This general reference provides an excellent overview of the Earth sciences. Chapter 10 is an extensive discussion of the Earth's crust, including useful illustrations and diagrams. Contains a glossary, an index, and recommendations for further reading.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall,

2008. This college-level textbook for introductory geology courses is well written and illustrated. There are very good chapters on the Earth's interior and plate tectonics.

Taylor, Stuart R., and Scott M. McLennan. *The Continental Crust: Its Composition and Evolution*. Boston: Blackwell Scientific, 1985. A text aimed at undergraduate and graduate geology and geophysics students as well as general Earth scientists. It is clearly written and has excellent, well-rounded scientific references.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996. Rigorously covers geophysics research and theories since 1960.

See also: Earth's Age; Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Differentiation; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Structure; Terrestrial Planets.

Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity

Category: Earth

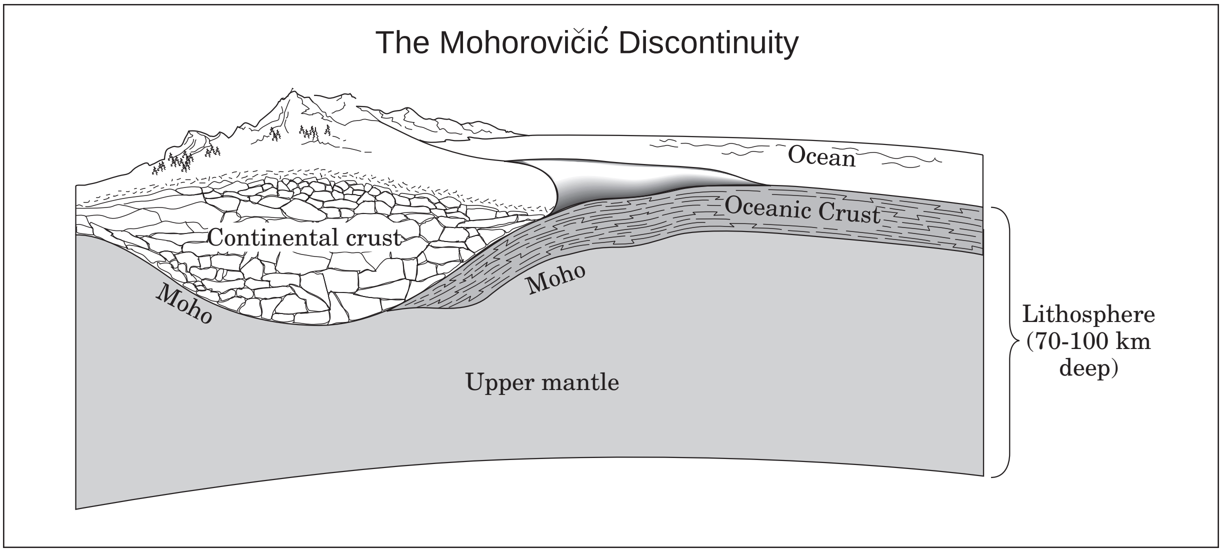
The Mohorovičić discontinuity, or Moho (the boundary between the crust and mantle), was discovered in 1909 through the observation of an abrupt change in the speed of seismic waves traveling below the surface of the Earth. Its discovery was among the first evidence for the now-famous "onion" model of the layers of the earth, but many of its fundamental properties are still not well understood.

OVERVIEW

On October 8, 1909, Central Europe was struck by a large earthquake centered in Croatia, near the village of Papuspsko. Andrija Mohorovičić, director of the Meteorological Observatory at the nearby University of Zagreb, collected data on the quake and its aftershocks. In the seismic records from stations at intermediate distances from the epicenter of the quake, he identified two sets of seismic P (longitudinal, or compressional) and S (transverse, or shear) waves, one set arriving at the recording stations sooner than the second. He correctly interpreted the first set of seismic waves to arrive as having traveled deeper and hence faster through material that had a higher density (and possibly different composition) than the crust material through which the shallower and slower second set of waves had passed. Just as light waves moving from water or glass into air bend or refract because of a sudden increase in speed, so also these seismic waves changed speed and refracted at some boundary, then traveled just under the boundary between these layers for some distance before being refracted again back into the lower-density layer and traveling through it to the surface. This boundary, now known to divide the crust from the mantle, was later named after its discoverer, but the somewhat unwieldy title Mohorovičić discontinuity commonly is shortened to Moho. Similar seismic boundaries have since been discovered between the mantle and outer core, and between the inner and outer cores.

The thickness of the crust, and hence the depth of the Moho, varies from 3 kilometers at mid-ocean ridges to 70 kilometers under young, high mountain ranges created by continental collisions, such as the Himalayas. It is important to note that the Moho does not mark the bottom of the tectonic plates; each plate is made of the crust and the uppermost mantle welded together to form the rigid lithosphere, which moves over the mantle's plastic layer known as the asthenosphere.

Since its discovery, the Moho has been the subject of intense scrutiny by geophysicists because it is the closest boundary to the Earth's surface and hence the easiest to study. One of the first debates was whether the Moho represented a transition between rock layers of different chemical composition or between rock layers of similar composition but different mineral and crystalline structure (phase changes caused by changes in temperature and pressure). Since the 1960's, the former model has been that favored by most geologists, although



some alternate models based on phase transitions still exist. The crust appears to be composed mainly of basaltic rocks in ocean floors and granitic rocks in continents, while the upper mantle is largely made of peridotite. This model is supported by the study and comparison of continental rocks, ophiolite complexes (which are slices of former ocean floor thrust above sea level), and xenoliths (which are pieces of the lower crust and upper mantle brought up to the surface by volcanic activity).

Worldwide studies of seismic data have added considerably to our understanding of the Moho. Seismic refraction data have been gathered not only from earthquakes but also from nuclear weapons testing and other underground explosions. Seismic reflection methods—which generate seismic waves by means of a near-surface explosion or truck-mounted vibrator—have also been widely used. These waves travel down through the crust, reflect off the Moho, and travel back to the surface, like light waves bouncing off a mirror. The out-and-back travel time can be used to determine the depth of the boundary, similar to the way a bat emits sounds that echo off obstacles in its path. Studies at some locations have also detected a change in the electrical conductivity of rocks near the Moho.

While this broad array of experimental data would suggest that the Moho is well defined, the opposite is actually the case. The problem is that the transition depths determined by these different methods often disagree. Therefore, the classic or seismic Moho (defined as the refraction transition, where the velocity of P waves jumps to 7.6 kilometers per second) may differ from the reflection boundary, the electrical conductivity boundary, and most important, from the rock-type or petrologic boundary, which is usually regarded as the true crust-mantle boundary. Compounding this problem is the fact that the seismic Moho is absent in some locations, with the velocity of the P waves only gradually increasing with depth. Taken together, these studies led to a questioning of the early model of the Moho as a thin, sharp, well-defined boundary in favor of a transition layer of definite thickness (perhaps several kilometers), over which the composition of the rocks changes. The Moho also has a complex, multilayered structure in regions of complex tectonic history. These observations have led geologists to view the Moho as a dynamic entity that evolves over time, in terms of both its petrology and its physical structure and substructures.

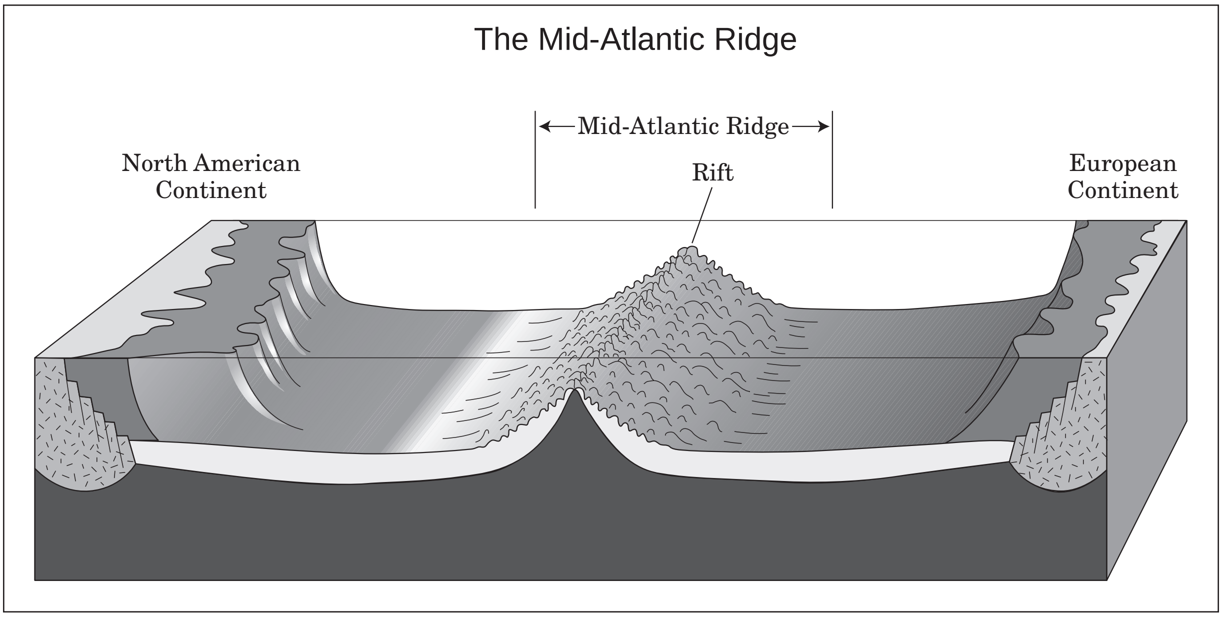
KNOWLEDGE GAINED

Although these generalizations about the nature of the Moho are fairly well established, the details of the structure and creation of the Moho are less well understood. Due to the difference in ocean crust and continental crust in terms of composition, thickness, and tectonic interactions, detailed studies of the Moho are usually divided into either oceanic or continental. In addition, the problem of understanding these fundamental processes is so complex that many researchers in this area utilize seismic and petrological data to model the Moho in a single geographic region rather than attempt to make a unified model of the entire Moho.

More is known about the oceanic Moho than its continental counterpart, largely thanks to studies of ophiolite complexes. Distributed worldwide, these samples of former (ancient) oceanic crust and upper mantle allow for direct study of the chemical and mineral composition of the rocks. Coupled with seismic refraction and reflection data, they paint a reasonably clear picture of the oceanic Moho and its history. Its structure is complex, with interlocking layers ranging in composition from mafic (basaltic) at the top to ultramafic at the bottom, finally merging with the peridotite of the mantle. The total thickness of the oceanic Moho layer ranges from 0 to 3 kilometers. Oceanic Moho is created from the same source as oceanic crust, namely the upwelling of magma at mid-ocean ridges such as the Mid-Atlantic Ridge (which is currently increasing the width of the Atlantic Ocean). It forms rather quickly in geologic terms, within a few thousand years, and after initial creation it is not significantly modified.

The continental Moho is much less well understood. Petrologic information can be gathered from xenoliths and exposed samples of the crust-mantle boundary uplifted by tectonic forces, although the latter must be used with caution, since such rocks have been changed by the same forces that lifted them to the surface. These tectonic forces, such as continental collisions, along with the much longer time frame covered by continental crust samples, lead to a greater complexity in the data. Because of the greater age of the continental rocks, the problem of how the Moho forms in continental regions is central to our understanding of how the early Earth differentiated into layers in the first place.

At least four hypotheses have been suggested for how the continental Moho forms. In the relic Moho model, the continental Moho is the relic of the oceanic Moho, surviving the assembly processes (such as continental collisions and accretion of terranes along the coast) that build the continents. How this might occur without severely disrupting the Moho is unclear, especially given the fact that the crust itself is certainly changed in this process. The magmatic



underplating hypothesis suggests that as continents assemble, new Moho material is added beneath them in a process similar to that which creates sills, horizontal intrusions of magma sandwiched between rock layers. In this model, the Moho would be younger than the continental crust, since it is created after the crustal material is set into place.

These first two hypotheses rely on igneous processes, while the final two hypotheses suggest the continental Moho consists of "reworked" rocks. In the metamorphic/metasomatic front model, the Moho is created by additional metamorphism of rocks of the lower crust and/or the upper mantle. Finally, the regional décollement hypothesis posits that the Moho forms as the crust and mantle physically decouple at structurally weak zones, especially under high temperatures.

It should be noted that this list is not suggested as being complete, and all four processes may occur in different geographical areas, depending on the particular geological conditions present. This conclusion has been suggested by researchers working on the Canadian LITHOPROBE program as a result of their study of the subsurface geology of North America. Therefore, it may be that there is no single explanation for the creation and evolution of the continental Moho, and the geologic history of each geographical region must be interpreted to develop a sensible model of its Moho structure and evolution.

CONTEXT

All methods described so far to study the Moho have relied on indirect testing, with the exception of samples of ophiolites and xenoliths. However, as previously mentioned, relying on these samples as truly representative of the Moho in general is unwise, since they may have been severely modified by tectonic forces. The most reliable rock sample would obviously be one obtained directly from the current Moho through drilling. However, given the great depths involved and the resulting technological difficulties, no direct Moho sample has yet been obtained, despite a number of programs designed to do so.

The most infamous such project was dubbed Project Mohole, proposed to the National Science Foundation in 1957 and funded until 1966. This three-stage project completed only its first phase, experimental drilling off the coasts of California and Mexico, before it was canceled by Congress (after a series of bureaucratic and financial problems). Despite the failure of the project, geologists remained committed to large-scale drilling projects in the ocean floor, leading to a series of ongoing projects such as the Deep Sea Drilling Project and Ocean Drilling Project. Although individual projects have reached depths of more than a kilometer, identifiable Moho samples have yet to be obtained, in keeping with the problems of identifying the seismic Moho with the petrological crust-mantle boundary and the complex nature of the Moho in general.

Therefore, nearly a century after the discovery of the Moho, its secrets continue to intrigue geologists. Despite the gaps in our understanding of the terrestrial crust-mantle boundary, planetary geologists are currently applying terrestrial methods and models to its neighbors in space. For example, lunar data suggest that the Moon's Moho lies about 30 to 70 kilometers beneath its surface, being shallowest beneath the maria. Models of Martian structure suggest that its Moho may lie between 6 and 100 kilometers beneath its surface. Such studies add to our understanding of the geologic evolution of these rocky worlds and how their evolution is similar to, yet differs from, that of Earth.

Kristine Larsen

FURTHER READING

Bascom, Willard. *A Hole in the Bottom of the Sea*. Garden City, N.Y.: Doubleday, 1961. An enthusiastic and detailed (but one-sided) popular-level survey of the scientific motivation behind Project Mohole, written by the project's original director.

Cook, Frederick A. "Fine Structure of the Continental Reflection Moho." *Geological Society of America Bulletin* 114, no. 1 (2002): 64-79. This technical review of seismic reflection data focuses on Canada yet contains a valuable overview of the topic and a lengthy bibliography.

Eaton, David W. "Multi-genetic Origin of the Continental Moho: Insights from LITHO-

PROBE." *Terra Nova* 18, no. 1 (2008): 34-43. This technical paper reviews the four hypotheses for the formation of the continental Moho described above. The bibliography is extensive.

Fowler, C. M. R. *The Solid Earth: An Introduction to Global Geophysics*. 2d ed. New York: Cambridge University Press, 2004. An updated version of a widely used textbook for introductory geophysics courses. Designed for students with some knowledge of physics and calculus.

Griffin, W. L., and Suzanne Y. O'Reilly. "Is the Continental Moho the Crust-Mantle Boundary?" *Geology* 15 (1987): 241-244. This seminal technical paper was among the first to succinctly and precisely challenge the assumption that the seismic and geochemical boundaries between the crust and mantle are identical.

Jarchow, Craig M., and George A. Thompson. "The Nature of the Mohorovičić Discontinuity." *Annual Review of Earth and Planetary Sciences* 17 (May, 1989): 475-506. A detailed overview of the general structure of the Moho and the main difficulties with the simplistic standard model.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. There is a very good chapter on the Earth's interior.

Van der Pluijm, Ben, and Stephen Marshak. *Earth's Structure*. 2d ed. New York: W. W. Norton, 2003. An introductory text on structural geology and tectonics. Designed for undergraduate students.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996. Covers geophysics research and theories since 1960. Includes information about Pangaea, the supercontinent cycle, and the reversals of the Earth's magnetic field. For general audiences.

See also: Earth-Moon Relations; Earth-Sun Relations; Earth System Science; Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Rotation; Earth's Shape; Earth's Structure; Planetary Interiors.

Earth's Differentiation

Category: Earth

Earth's differentiation refers to the separation of matter in the early Earth into "layers" or zones with different chemical compositions and physical properties, during or shortly after its formation about 4.5 to 4.6 billion years ago. The structure that developed consists of a metallic core surrounded by a rocky mantle overlain by a thin, rocky crust.

OVERVIEW

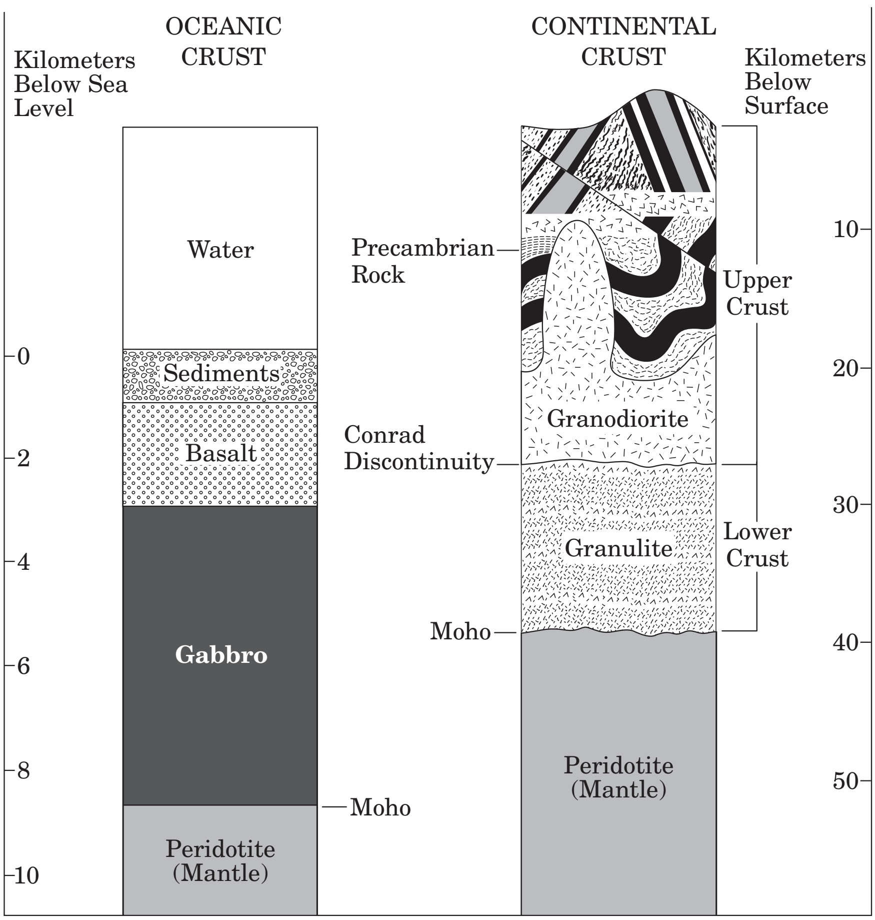
The Earth consists of nearly spherical concentric layers arranged according to density. The core, composed mainly of metals like iron and nickel, is the central part of the Earth. Its radius is about 3,400 kilometers, and its density ranges between about 10 to 13 grams per cubic centimeter. The crust is the outermost layer at the surface of the Earth. This layer is very thin, its thickness ranging from as little 5 kilometers under ocean basins to as much as 70 kilometers under young, high mountain ranges. The average density of continental crust is about 2.7 grams per cubic centimeter, while ocean basin crust is somewhat denser, at about 3.0 grams per cubic centimeter.

The most common minerals in the crust are the silicates, which have as their basic crystal lattice structure the silicon-oxygen tetrahedron (one silicon atom surrounded three-dimensionally by four evenly spaced oxygen atoms). These tetrahedra combine with atoms of many other elements, especially aluminum, iron, calcium, sodium, potassium, and magnesium, to form the large class of minerals called silicates. Between the crust and the core is the mantle, which is rocky like the crust, but composed of denser silicate minerals with larger percentages of iron and magnesium. The mantle is about 2,900 kilometers thick, and ranges in density from about 3.3 grams per cubic centimeter at its top, just below the crust, to about 5.6 grams per cubic centimeter at its base, just above the core.

Earth and the other planets in the solar system formed about the same time the Sun did, approximately 4.5 to 4.6 billion years ago. Planet formation occurred by condensation of solid dust grains in the solar nebula, followed by accretion of those solid grains to form planetesimals that grew in size through continued accretion. There are two major models of how this condensation and accretion occurred: homogeneous accretion and inhomogeneous or heterogenous accretion.

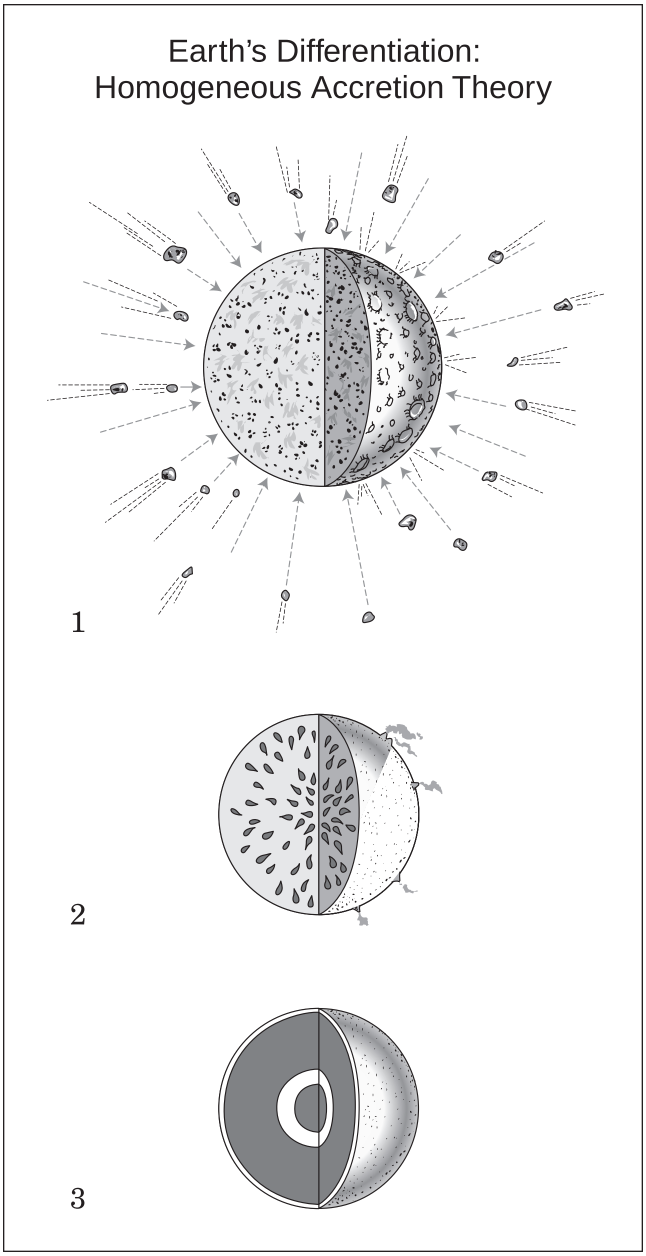
If cooling of the solar nebula and the resulting condensation occurred faster than accretion, the material swept up in the accretion process would have been well mixed in its composition. The early Earth would have been homogeneous, or undifferentiated, composed of random grains of metal alloys, oxides, silicates, and carbonaceous materials throughout. During or shortly after accretion, the early Earth became hot enough to at least partially melt.

Comparison of Zones of Oceanic and Continental Crust



Dense blobs of molten metals like iron and nickel sank to the center to form the core, while less dense silicate and oxide minerals remained to form the mantle. The material near the surface also melted, becoming a magma ocean a few hundred kilometers deep. When it cooled and hardened, it formed a thin crust of basaltic silicate minerals probably similar to today’s ocean basin crust. Radioactive elements such as uranium and thorium are relatively abundant in the crust because of their tendency to combine with low-density crustal minerals.

In order for this homogeneous accretion model to be correct, there must have been some processes that heated the early Earth enough to at least partially melt. Several likely processes have been proposed. First, radioactive isotopes were much more common in the early Earth because they had not yet had time to decay; radioactive decay of these isotopes would have produced much heat. This heat source has greatly decreased through time as the isotopes with short half-lives have decayed, but it is still probably the major source maintaining Earth’s internal heat today. Second, as planetesimals continued to collide with the proto-Earth, their kinetic energy (energy of motion) was converted on impact into heat energy. Third, heat is produced by gravitational compression. As more and more material was added to Earth, progressively greater temperatures were generated in the interior. Furthermore, as Earth grew larger in size, it tended to insulate itself, raising the internal temperature by trapping interior heat and impeding its outward flow. The homogeneous accretion model requires that the combined effects of all these process raised the interior temperature of Earth enough to cause at least partial melting. Once that occurred, dense lumps of metal began to sink toward the center. In doing so, they converted gravitational potential energy into more heat, leading to more melting, more sinking, even more heating, and so on, quickly becoming a



Homogeneous accretion theory states that the Earth formed by randomly sweeping up debris (meteors, dust) in its orbit around the Sun. (1) This process, which occurred about 4.6 billion years ago, caused the debris to be added on (accreted) to the early Earth in a sort of “snowball” effect that made the Earth progressively larger and still undifferentiated, or homogeneous. (2) Gravitational compression, increasing temperatures in the interior, initiated differentiation, as iron, nickel, and other denser elements sank to the core. (3) The result was that lighter elements rose to the surface to form the crust, and elements of intermediate density stayed between, forming today’s core, mantle, and crust.

runaway collapse of a metallic core. So much heat would have been produced by the collapse that the entire core and probably much of the mantle would have been molten.

However, some computer models of the early Earth indicate that not enough heat would have been generated internally to result in partial melting. An alternative to this possible heating problem of homogeneous accretion is provided by the heterogeneous or inhomogeneous accretion model. If accretion occurred faster than cooling of the solar nebula and the resulting condensation, then the material swept up in the accretion process would have varied with time. The first to condense and accrete would be the materials with higher melting points, like metal alloys and oxides, forming metal-rich cores. Later as the temperature in the solar nebula dropped further, silicate minerals would condense and be swept up, forming rocky mantles around the metallic cores, and eventually even less dense rocky crusts at the surface.

In this heterogeneous or inhomogeneous accretion model, differentiation occurred during accretion as a result of changes in what had condensed and was ready to be picked up. Instead of accreting as an originally homogeneous body, which then became layered as denser material sank and less dense material rose through melting, Earth would have accreted its layers in sequence.

In spite of their differences, the two models agree on a number of major points about Earth's formation, the most basic being that the Earth originated from condensation and accretion in a large cloud of dust and debris around the early Sun about 4.5 to 4.6 billion years ago. The models also agree that differentiation began very early in Earth's history, either as part of the accretion process (in the inhomogeneous accretion model) or within the first few tens of millions of years (in the homogeneous accretion model). A primitive crust must have developed within the first few hundred million years. Zircon crystals possibly dating back to 4.4 billion years ago have been found in somewhat younger rocks in Australia's Jack Hills region, and the oldest known definitely continental rocks—the Acasta gneiss from Canada's Northwest Territories—date back to 4 billion years ago.

Whichever of the two models is correct, the differentiation of the Earth did not end with the formation of the core, mantle, and crust. The early crust almost certainly consisted mainly of mafic and ultramafic rocks, which are composed of relatively dense silicate minerals rich in iron and magnesium. Today, there are two distinct types of crust: continental crust, composed mainly of granitic or felsic rocks with an average density of about 2.7 grams per cubic centimeter, and oceanic crust, composed mostly of basaltic or mafic rocks with an average density of about 3.0 grams per cubic centimeter. The differentiation of crustal material is thought to be the result of plate tectonic processes that presumably have been occurring ever since the first primitive crust was formed. The motion of the tectonic plates and the crustal differentiation that accompanies it are thought to be driven by slow convection currents in the mantle.

Rocks from oceanic crust underlying the ocean basins are mainly basaltic or mafic, consisting mostly of minerals that are not quite as rich in iron and magnesium and thus not as dense as ultramafic rocks, although they are denser than granitic rocks. The separation of mafic from ultramafic material presumably takes place at the ocean ridge-rift system located along divergent plate boundaries. As the tectonic plates spread apart, mafic material rises to produce new basaltic oceanic crust while the ultramafic material remains in the mantle.

Rocks from continental crust are mainly granitic or felsic, consisting primarily of less dense silicate minerals rich in potassium and sodium. The separation of granitic continental rocks from basaltic oceanic crust is thought to be the result of subduction, which occurs at convergent plate boundaries, where an oceanic plate descends (or subducts) under another plate, either oceanic or continental. As the subducting oceanic plate descends back into the mantle, it is heated and differentially melted. The minerals that preferentially melt are those with lower melting temperatures. These are mostly the lower density silicate minerals common in granitic rocks. They rise as molten blobs of magma and produce the granitic rocks of continental crust, either erupting through the surface as volcanoes, or cooling and hardening below the surface in magma chambers. The bulk of the oceanic crust, consisting of mafic rocks, returns into the mantle.

The early Earth along with the other planets was heavily bombarded by planetesimals in the latter stages of accretion, and this bombardment may have contributed to the formation of early continents. Large impacts would have melted the early crust and upper mantle. Denser minerals would have sunk while less dense minerals would have risen during cooling and crystallization, possibly forming a granitic continental "nucleus" around which additional continental material was added later by subduction.

Whether large impacts contributed to early continent formation or not, the amount of continental crust has continued to increase with time, as new continental crust is produced at subduction zones, and once formed, generally remains continental crust (although it can change from one rock type to another). In contrast, oceanic crust is continually recycled, being produced at the ocean ridge/rift system and consumed at subduction zones.

METHODS OF STUDY

It is not possible to study directly the differentiation of the Earth because it occurred at least 4 billion years ago. Most of our current theories and models are based on inferences from indirect methods of study. Some tentative clues about the first stages of condensation and accretion are provided by the protoplanetary disks (or proplyds) that have been discovered around very young stars.

Most meteorites are remnants from the early solar system and thus yield more evidence about the condensation and accretion processes that occurred then. Meteorites are divided into three main groups based on composition: stony meteorites, iron meteorites, and stony-iron meteorites. A subgroup of stony meteorites are the chondrites, which consist of glassy spherules called chondrules embedded in a mixture of small grains of various silicate minerals and metal alloys and sometimes carbonaceous compounds. These meteorites are thought to be relatively unaltered samples of the accretion process. Another subgroup of stony meteorites are the achondrites, which lack chondrules and instead show evidence of heating and melting followed by cooling and crystallization. The iron meteorites, composed mostly of iron and nickel, appear to have formed under extremely high temperatures and pressures such as exist in planetary cores. The achondrites and the iron meteorites are thought to have originated from early planetesimals or protoplanets that underwent differentiation into rocky mantles and metallic cores before they were shattered by collisions in the early solar system. The stony-iron meteorites, consisting of intermixed rocky and metallic parts, may be pieces of planetesimals or protoplanets that only partially differentiated before they were broken up by collisions.

Laboratory studies used to model differentiation show that iron and nickel will separate from silicate minerals at an early stage, and thus would be available to form a metallic core.

The next major group to separate are the silicate minerals olivine and pyroxene; rich in iron and magnesium, they form the ultramafic rock peridotite, which is thought to be a major constituent of the mantle. Laboratory studies of peridotite also indicate how oceanic and continental crust would differentiate from it. As peridotite rises up from the mantle at divergent plate boundaries, the pressure and temperature of the magma's environment begin to decrease dramatically. Ultramafic material remains in the mantle, while mafic material rises to become new basaltic oceanic crust. At convergent plate boundaries where subduction is occurring, the basaltic oceanic crust is pushed underneath the other plate back into the mantle, where it heats and partially melts. Laboratory studies show that its partial melting creates magmas that rise, cool, and harden into minerals found in granitic continental crust.

The structure and physical properties of Earth's interior are revealed by seismic waves from earthquakes that pass through Earth. Two types of body waves that travel through the body or interior of Earth are generated. P waves (or primary waves, so named because they travel faster and arrive first) oscillate back and forth in the direction of travel and will go through solid or liquid material. S waves (or secondary waves, so named because they travel more slowly and arrive second) oscillate at right angles to the direction of travel and will go only through solids but not through liquids. S waves are not received at locations on the opposite side of Earth from an earthquake. This "S-wave shadow zone" indicates that Earth has a core that is composed of molten material. On the other hand, P waves are received on the opposite side of the Earth, but they arrive sooner than they would if the entire core were molten, indicating that the inner part of the core is solid, while only the outer core is molten. Using records of the arrival times of P and S waves at seismic stations around the globe, scientists can determine the speeds of P and S waves as they travel through different parts of Earth's interior, and the speeds (together with knowledge of whether the material is solid or molten) reveal that the density increases with depth. Studying how and where the seismic waves are reflected or refracted indicates the depth of the boundaries between the crust and mantle, the mantle and outer core, and the outer core and inner core.

CONTEXT

The events and processes involved in Earth's differentiation shaped the Earth we now inhabit. The better those events and processes are understood, the better we will understand the planet as it is today.

Ores are rocks that are naturally enriched in some desirable mineral resource through differentiation. Many common rocks contain trace amounts of usable minerals, but they are present in such small quantities that it is not economical to try to extract them. Knowledge of differentiation processes can aid in locating ore bodies that are economical to mine.

Most earthquakes are produced by plate motion. The processes that formed the mantle and crust established the pattern of plate tectonics. Knowledge of plate tectonics and the differentiation occurring as part of it adds to our knowledge of how and where earthquakes occur.

Studies of Earth's differentiation can be extended to better understand the formation and someday possibly to utilize the resources of other solar-system objects. The other terrestrial planets—Mercury, Venus, and Mars, and also Earth's moon—underwent planetary differentiation at about the same time as did Earth. The asteroids are remnants of the planetesimals that never formed large planets. All these objects probably are rich in useful materials that we may someday have the technology to retrieve.

Earth's differentiation is not yet complete. The crust continues to evolve due to plate tectonics. Physiochemical changes continue to occur in the core and mantle as the interior slowly cools. Extrapolating these processes and comparing Earth to planets that have or have not undergone similar processes makes it possible to draw conclusions regarding the long-term future of Earth.

Michael L. McKinney

FURTHER READING

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that have investigated all planets with the exception of Pluto. Compares Earth to the other planets.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college textbook beyond the introductory level, its approach is based on comparative planetology. Provides much information on condensation and accretion in the solar nebula, as well as differentiation in the terrestrial planets generally and Earth in particular.

Levin, Harold L. *The Earth Through Time*. 5th ed. Fort Worth, Tex.: Saunders College Publishing, 1996. A highly respected and widely used college textbook for introductory historical geology. Very well illustrated and clearly written; an excellent introduction to the subject. Technical references for further study. Contains a summary of Earth's growth and differentiation.

Ozima, Minoru. *The Earth: Its Birth and Growth*. Translated by J. F. Wakabayashi. New York: Cambridge University Press, 1981. An excellent overview of Earth's differentiation from the beginning of planetary condensation to the present. Suitable for interested laypersons and advanced high school students. Technical in parts, but covers many basic concepts.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. It has a full chapter on the origin and historical development of Earth that includes differentiation processes.

Wicander, R., and J. Monroe. *Historical Geology*. St. Paul, Minn.: West, 2006. A survey of Earth history, with a good summary discussion of Earth's differentiation. A basic college-level text, but readable for the layperson and the advanced high school student.

See also: Earth-Moon Relations; Earth-Sun Relations; Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Rotation; Earth's Shape; Earth's Structure; Lunar History; Solar System: Element Distribution.

Earth's Magnetic Field: Origins

Category: Earth

The Earth has a dipole magnetic field that is roughly aligned with its rotational axis. A dynamo effect in the Earth's molten outer core is the most likely source of most of the magnetic field.

OVERVIEW

The Earth's magnetic field is primarily a dipole field, meaning it has two well-defined magnetic poles, called "north" and "south" (the prefix "di" comes from a Greek term for "two"). This is the type of field produced by a bar magnet or an electric current flowing in a wire loop. The Earth's iron core once was thought to act like a giant bar magnet, possibly due to a remanent field frozen in place from some primordial magnetic field that existed when our solar system was forming. Now, however, electrical currents produced by fluid motions in the molten outer part of the Earth's iron core are theorized to be the source of the magnetic field. This conclusion is based on models of the Earth's interior structure and variations in the magnetic field over both historic and geologic timescales.

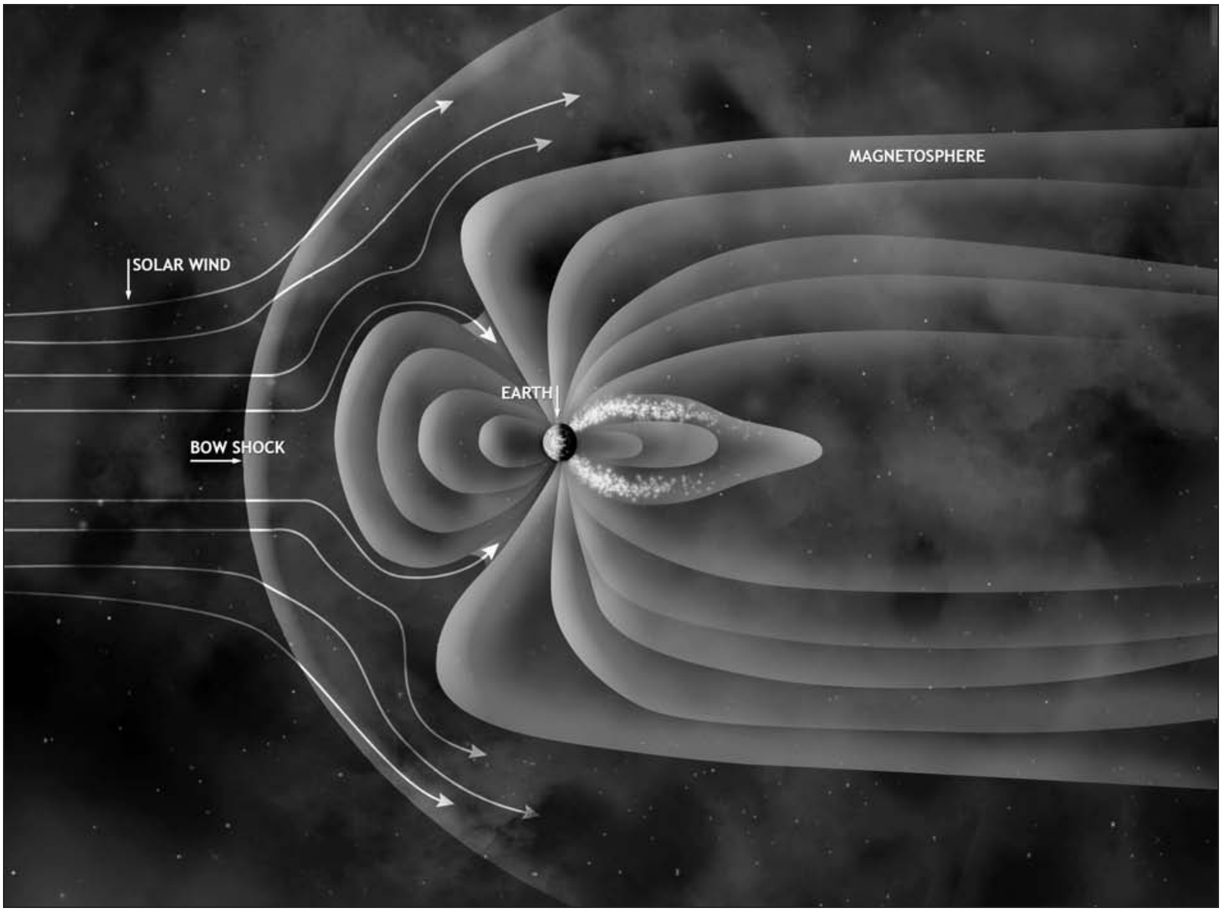
The ultimate source of any magnetic field is the movement of electric charges. Wires carrying electric currents, for example, have magnetic fields around them because of the electric charges (electrons) moving through the wires. The electrons surrounding the nucleus of an atom are moving, and this produces a minute magnetic field. In a magnet, the atoms are aligned in such a way that these small fields add together to produce the larger field of the magnet. The conclusion, therefore, is that electric currents within the Earth produce its magnetic field through a process referred to as the geodynamo.

To determine the Earth's interior structure, seismic waves from earthquakes act as probes as they pass through the Earth. They reveal that the Earth's interior consists of three major zones or layers: the crust, mantle, and core. The crust and underlying mantle are composed mainly of rocky material, which is a good electrical insulator. The innermost region, the core, is composed of metals, most probably iron with a small percentage of nickel and an even smaller percentage of other elements, and thus it is a good electrical conductor. The inner core (out to a radius of about 1,300 kilometers) is solid, but the outer core (from 1,300 to about 3,500 kilometers radius) is molten.

The temperature of the solid inner core is estimated to be about 5,800 kelvins. Heat flowing outward through the outer molten core sets up convection currents, in which hotter, less dense fluid rises. When it transfers its heat to the overlying mantle, the fluid cools, becomes denser, and sinks. The convection may also be partly driven chemically. If iron crystallizes at the bottom of the molten outer core to add to the solid inner core, the remaining fluid contains less iron and so is less dense, augmenting the thermally driven upward motion. Simple convection currents are deflected by the Earth's rotation in a process called the Coriolis effect. Computer models of the outer molten core that incorporate both convection and rotation show that the fluid moves in a number of spiraling columns aligned roughly parallel with the Earth's rotational axis.

All that is needed to "jump start" the geodynamo is a weak background magnetic field, perhaps provided by the solar wind or a remanent primordial field. As the metallic fluid moves through the background field, electrical currents are induced that in turn generate their own magnetic fields. This produces a positive feedback that reinforces the electrical currents and the overall magnetic field. The approximate alignment of the fluid's spiral motion with the rotational axis produces a dipole field with the magnetic poles near the rotational poles. The energy to produce the stronger magnetic field comes from the motion of convection and rotation. This process does not continue generating an ever-increasing field; it levels off since it becomes harder to generate an even stronger field as the field strength increases.

The geodynamo process explains many features of the Earth's magnetic field. Although the magnetic poles apparently remain close to the rotational poles, there is a shift in the position of the magnetic poles and changes in the



When the solar wind meets the Earth's magnetic field, a "bow shock" effect results, and the hot gases and radiation of the solar wind are deflected. Electrons are channeled by the magnetic field to create auroras during periods of high solar activity. (NASA/CXC/M. Weiss) field strength over periods of years to centuries; this could be caused by changes in the convection currents within the molten outer core.

Furthermore, paleomagnetic studies indicate that the Earth's magnetic field has reversed polarity at irregular intervals many times in the geologic past, the last reversal occurring about 700,000 years ago. The geodynamo process is unstable over long periods of time and can decay and regrow with changed polarity. Geologists have constructed models of dynamos that are simple versions of the geodynamo, and when set in operation, these models display changes in the field's intensity and polarity. The geodynamo explains the origin of about 90 to 95 percent of the Earth's magnetic field; the rest probably comes from fields associated with magnetic minerals in the Earth's crust, more complicated irregularities in the convective motions of the molten outer core, and external sources such as the solar wind interacting with the Earth's ionosphere.

METHODS OF STUDY

The Earth's magnetic field is generated in its interior. A clue to the composition of the interior is provided by the Earth's average density. Dividing the Earth's mass by its volume shows the average density to be about 5.5 times the density of water. Common rocks from the surface are about 3 times denser than water. Therefore a portion of the Earth's interior must be much denser than surface rocks in order to yield the average value. Only metals have the required density, but some metals, such as aluminum, are too low in density, and others, such as uranium, are too high. Still others, such as gold and silver, are close to the required density but are too rare. Iron is a good candidate, since it has the right density and is fairly abundant.

The Earth's interior structure can be probed using seismic waves produced by earthquakes. Body waves from the earthquake travel through the Earth's interior. Their speed and direction of travel are determined by the density and elastic properties of the material through which they are traveling. The waves also are reflected off boundaries between different layers. When the transmitted and reflected waves reach the surface, they are recorded on seismographs. Analysis of the seismograms obtained at seismic stations all around the globe reveals that the Earth has three main zones or layers: the surface crust, the mantle, and the central core. One type of body waves, called S waves, are transverse waves that can travel only through solids, not liquids. Their presence or absence recorded on seismograms reveals which parts of the interior are solid or liquid. The crust and mantle are solid (except for isolated pockets of molten material called magma). The inner part of the core also is solid, but the outer core is molten.

Models of the Earth's interior that combine chemical composition estimates with seismic-wave data indicate that the crust and mantle are composed of rocky material, while the core (both molten outer and solid inner parts) is mostly iron, with some nickel and other elements. The composition and pressure from these models can be used to calculate the melting-point temperature as a function of depth. The mechanical properties of the layers (whether solid or liquid, as indicated by seismic waves) can then be used to determine whether the actual temperature is below or above the calculated melting-point temperature. Anchored by measurements of heat flow at the surface and the increase of temperature with depth recorded in mines and wells, the geotherm—a graph of actual temperature versus depth—can be drawn. This is how the temperature of the Earth's core, about 5,800 kelvins, is determined, showing that there is enough heat energy to drive the convection necessary for the geodynamo process.

Remanent magnetism, the evidence of past magnetic fields preserved in some igneous and sedimentary rocks, can be measured with various types of magnetometers. These data, together with records of changes in the Earth's magnetic field during historic times, show how the field varies in strength and orientation and has even reversed polarity many times in the past.

CONTEXT

The geodynamo mechanism that generates the Earth's magnetic field also can explain the presence or absence of magnetic fields for other planets, their moons, and the Sun. Mars and Earth's moon have extremely weak magnetic fields, probably because their iron cores are so small, and they may have cooled to the point that their iron cores are no longer molten so convection cannot occur. Venus also has an extremely weak field; although it is nearly the same size and mass as Earth and probably has a similar internal structure with a molten outer core, it rotates very slowly.

Mercury has a magnetic field about an order of magnitude stronger than the fields of Venus, Mars, and the Moon, but about two orders of magnitude weaker than the field of the Earth; although Mercury rotates slowly, its relatively large iron core might still have a molten convective zone. Jupiter's magnetic field is more than ten times stronger than Earth's, and Saturn's is about two-thirds as strong as Earth's; their fields are probably due to convection occurring in the liquid metallic hydrogen in their interiors (Jupiter has much more than Saturn) and their rapid rotation. Jupiter's moons Europa and Ganymede also have small magnetic fields, probably due to convection in electrically conductive salty oceans beneath their icy crusts. The Sun has a strong magnetic field that reverses polarity every eleven years; it is produced by the movement of ionized gas in the convection zone of the Sun's interior.

Stephen J. Shulik

FURTHER READING

Busse, F. H. "Recent Developments in the Dynamo Theory of Planetary Magnetism." In *Annual Review of Earth and Planetary Sciences* 11 (May, 1983): 241-268. An outline of the dynamo theory is given, with models of the dynamo for various planets. Observational evidence is discussed, along with the paleomagnetic data, geomagnetic reversals, and secular variation. Includes references and figures.

Fowler, C. M. R. *The Solid Earth: An Introduction to Global Geophysics*. 2d ed. New York: Cambridge University Press, 2004. An updated version of a widely used textbook for introductory geophysics courses. Designed for students with some knowledge of physics and calculus.

Garland, G. D. *Introduction to Geophysics*. 2d ed. Philadelphia: W. B. Saunders, 1979. Used as a text for introductory geophysics, this book covers, in sections 17.4 and 17.5, the cause of the main field and the dynamo theory. A few equations, but many figures and graphs that are of interest to the less technically informed reader. At the end of the chapter is a listing of thirty-two references.

Gubbins, D., and T. G. Masters. "Driving Mechanisms for the Earth's Dynamo." In *Advances in Geophysics*. Vol. 21, edited by B. Saltzman. New York: Academic Press, 1979. This article looks at such topics as the physical and chemical properties of the core and energy sources for the magnetic field. References are located at the end of the article. Mathematics and numerous figures and tables are included.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college textbook beyond the introductory level, its approach is based on comparative planetology. Has a section on the generation of planetary magnetic fields by the dynamo mechanism.

Lapedes, D. N., ed. *McGraw-Hill Encyclopedia of Geological Sciences*. New York: McGraw-Hill, 1978. Pages 704-708, under the heading "Rock Magnetism," provide a concise description of many aspects associated with rock magnetism: how rock magnetization occurs, the present field, magnetic reversals, field generation, secular variation, and apparent polar wandering, among other subjects. Very readable, with no mathematics and a fair number of graphs, tables, and figures.

Merrill, R. T., and M. W. McElhinney. *The Earth's Magnetic Field*. New York: Academic Press, 1983. The authors cover much of the material associated with the Earth's field. Chapters 7 and 8 deal with the origin of the field, and chapter 9 covers the origin of secular variation and field reversals. Mathematical equations and thirty-eight pages of references. Numerous tables and figures.

Motz, L., ed. *Rediscovery of the Earth*. New York: Van Nostrand Reinhold, 1979. As a collection of articles for the nonscientist by scientists renowned in their respective fields, the text makes very interesting reading, augmented with many colorful illustrations. The chapter "The Earth's Magnetic Field and Its

Variations" is written by Dr. Takesi Nagata, who has authored hundreds of articles on diverse aspects of geophysics besides the Earth's magnetic field, and covers a wide range of magnetic field topics. Two pages are devoted to the origin of the field. Includes a small amount of mathematics and only a few references.

Smith, D. G., ed. *The Cambridge Encyclopedia of Earth Sciences*. New York: Crown, 1981. Chapter 7, "The Earth as a Magnet," contains a discussion of the field's origin. The text is well written at a nontechnical level, with many colorful diagrams and figures.

Stacey, F. D. *Physics of the Earth*. New York: John Wiley & Sons, 1977. Under section 8.4, "Generation of the Main Field," the author provides a short, technical description of the origin of the field, which will be of interest to the more advanced student. Equations are rather formidable, but several figures illustrating the dynamo effect are included. A large number of references at the end of the text. Many other areas of geophysics are covered at a technical level.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. The chapter on the Earth's interior has a section on the generation of the magnetic field by the geodynamo process. The chapter on plate tectonics has a section on geomagnetic reversals.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996. Covers geophysics research and theories since 1960. Includes information about Pangaea, the supercontinent cycle, and the reversals of the Earth's magnetic field. For general audiences.

See also: Auroras; Earth-Moon Relations; Earth-Sun Relations; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Planetary Magnetospheres; Planetary Orbits: Couplings and Resonances; Van Allen Radiation Belts.

Earth's Magnetic Field: Secular Variation

Category: Earth

At every point on the Earth, its magnetic field has a direction, as indicated by a compass needle free to pivot three-dimensionally, and an intensity or strength. The direction and intensity of the magnetic field change over timescales of years to millennia, a phenomenon known as secular variation. Over longer geologic timescales of tens of thousands to millions of years, the Earth's field reverses polarity, a phenomenon called geomagnetic reversal.

OVERVIEW

Secular variation of the Earth's magnetic field refers to changes in the field's direction and intensity, a phenomenon manifested everywhere on the Earth's surface. Its most obvious effect is a gradual shift in the direction which an ordinary compass needle points. It also is seen in changes of inclination, the angle at which a magnetic needle suspended by its center tilts below the horizontal, as well as variations in the field intensity or strength. These changes appear to be noncyclic and occur over timescales of years to millennia.

The direction of the Earth's magnetic field at any point is specified by two angles called declination and inclination. The north end of a magnetic compass needle points approximately to the north, but not exactly. The angle between geographic or true north (defined by the Earth's north rotational pole) and magnetic north originally was called magnetic variation, but now is called declination. It was first noticed around the twelfth century, when it was thought to be caused by abnormalities in the compass needle's magnetization or suspension. However, by the early sixteenth century, Europeans had accepted declination as a phenomenon of the Earth's magnetism. Inclination, the downward tilt of a compass needle free to pivot three-dimensionally, was also discovered during that century. Hence William Gilbert could write in 1600 of both declination and inclination as natural features of the Earth's magnetism.

By the early sixteenth century, Europeans had noticed that the declination varies from place to place (and this helped convince them that declination was a feature of the Earth's field and not due to flaws in their compasses). The discovery arose in the practices of navigation, chart making, and crafting of magnetic compasses—all activities connected with exploration. Perhaps Christopher Columbus, and certainly Sebastian Cabot, noted that while compass needles pointed east of north near Europe, they pointed west of north in the New World.

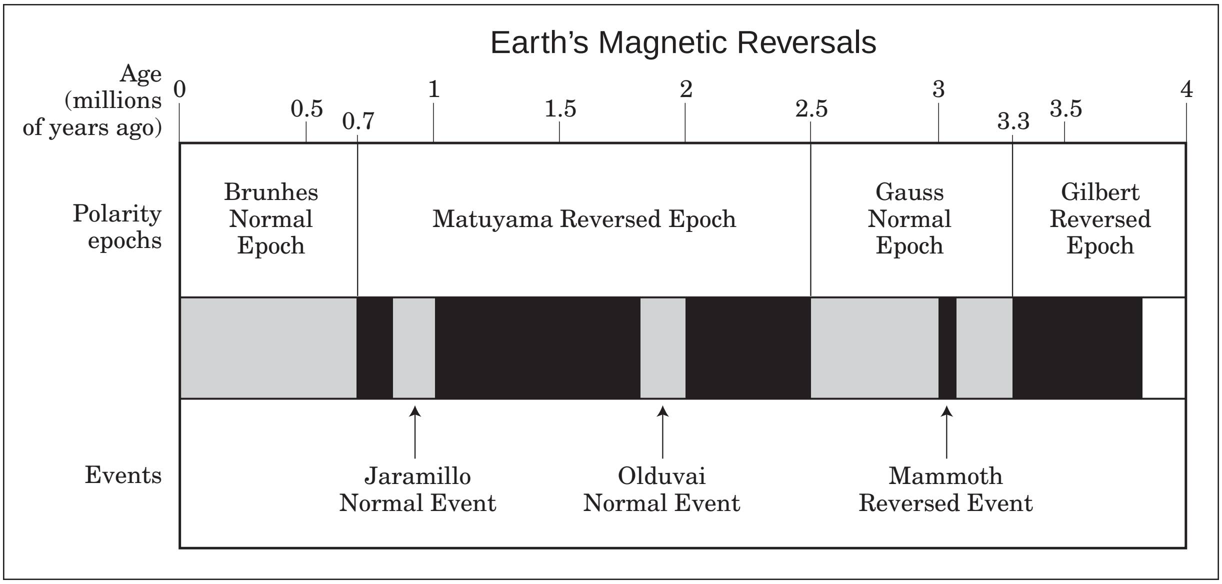
All three measures of the magnetic field—declination, inclination, and intensity—vary over the entire planet. These variations are most easily depicted with maps on which curved lines connect points that have the same value of one of these three parameters. For example, maps that display curved lines of equal magnetic declination are called isogonic maps. The first such printed map was produced in about 1701 by Edmond Halley of comet fame. Initially it was hoped that isogonic maps could be used to determine longitude by compass. Maps similar to isogonic maps but displaying lines of equal inclination or equal intensity also can be drawn.

Meanwhile, Henry Gellibrand announced in 1635 that declination changes over time as well as space. He found that the magnetic declination for London had shifted from 11.3° east of north in 1580 to 4.1° east of north in 1634. Later investigators discovered that the inclination and the intensity of the magnetic field also gradually change. For example, between 1700 and 1900, the inclination at London decreased from about 75° to 67°. Currently the overall intensity of the dipole field is decreasing at the rate of about 6 percent per century. If the field were to continue to decrease at this rate, it would drop to zero in about sixteen hundred years.

The agonic line is the "line of zero declination," along which compass needles point exactly toward geographic or true north. One aspect of secular variation has been the westward drift of the agonic line. This drift can be depicted on maps; just as one can map the magnetic parameters, one can also map how these parameters change, by drawing curved lines connecting points that change at the same rate. For example, all points where the declination is shifting westward at 10° per century would be connected together. These charts, known as isoporic charts, came into wide use in the mid-twentieth century. Areas of most rapid change are called isoporic foci. These isoporic foci are drifting westward, just as the agonic line is. While this westward drift has been a persistent feature of secular variation since Gellibrand, first pointed it out, some evidence exists for eastward drifts during prehistoric times.

Geologic evidence of ancient magnetic fields preserved in some igneous and sedimentary rocks shows that the geomagnetic field has reversed polarity many times over intervals of tens of thousands to millions of years. These geomagnetic reversals have played a major role in plate tectonics. They provide evidence of seafloor spreading and can be used to determine its rate. They indicate the locations of continents in the geologic past and thus can be used to trace continental drift.

The geomagnetic field and its secular changes traditionally have been attributed to the Earth's interior. Thus theories of the source of the field and the causes of its secular variation are necessarily indirect and extremely diverse, given the inaccessibility of the interior for direct study. About four hundred years ago, Gilbert suggested that the Earth behaves as if it had a bar magnet or magnetic dipole of extraordinary intensity at its center. In 1674, Robert Hooke asserted that the magnetic dipole axis of the Earth is tilted about 10° from the axis of rotation and that the dipole axis rotates westward around the rotational axis every 370 years. In 1683, Halley proposed that a double dipole pattern with four magnetic poles provided a better fit to worldwide declination data than just two magnetic poles, and in 1692, he suggested that his four poles could explain secular variation. Two of these poles he assigned to the Earth's outer crust and the other two to a central nucleus, which rotated slightly more slowly than the crust, on the same axis. The crustal magnetic poles were fixed in place, and as the nucleus, rotating a bit more slowly, drifted slowly westward relative to the crust, so did its magnetic poles. This explained, he thought, the drift of the agonic line. Theories that the core is per-



Geologic evidence of ancient magnetic fields preserved in some igneous and sedimentary rocks shows that Earth's geomagnetic field has reversed polarity many times over intervals of tens of thousands to millions of years.

manently magnetized were later ruled out when models of the Earth's interior showed that it is too hot; its temperature is above the Curie temperature of all known permanently magnetized materials. The Curie temperature (or Curie point) is the temperature above which a material is no longer permanently magnetic.

It is known that moving electrical charges generate magnetic fields. In particular, an electric current flowing around a wire loop produces a magnetic dipole field through the center of the loop. During the nineteenth and twentieth centuries, theories were developed that attributed the magnetic field and its secular variations to electric currents in the Earth's interior. One hypothesis was that the rotation of the Earth's iron core carried the charges with it, and this motion generated the field. This theory reached its highest state of development around 1950 in work by Patrick M. S. Blackett, but since then it has gradually lost favor. An alternative hypothesis is that the flow of molten metal in the interior carries the charge that generates the field. First introduced in rudimentary form in the nineteenth century, this theory became increasingly sophisticated with the investigations of Walter Elsasser beginning in 1939 and Sir Edward Crisp Bullard starting in 1948. Elsasser proposed that the combination of the movement of molten metal and the simultaneous flow of electricity in it produced both the Earth's main dipole field and its secular variations. This dynamo was driven, he suggested, by the heat generated by the decay of radioactive materials in the interior. Convection of hotter, less dense materials upward and of colder, denser materials downward, he said, produced the dynamo.

There now is general agreement that the geomagnetic field and its secular variation are the result of fluid motions in the molten outer part of the Earth's metallic core. Various models have been developed to show that convective motion in the molten outer core, modified by the Coriolis effect due to the Earth's rotation, can produce the observed field and its secular variations. Metals are good electrical conductors because electrons can move easily through metals. The molten metal of the outer core, as it moves through the magnetic field, makes electrons in the metal move, inducing electric currents in the metal that in turn generate the magnetic field. Thus the geodynamo is self-sustaining, but it is not a perpetual motion machine; it needs an energy source to drive the motion. The geodynamo does not create its magnetic field from nothing; rather, it converts some other form of energy into magnetic energy. The two most probable energy sources to drive the convection and generate the field are heat from the decay of radioactive materials and the crystallization and settling of iron (and other dense metals) to the solid inner core.

In the end, one must remember that models of the geodynamo and its energy sources are tentative. Many debates are still waged over the details of the geodynamo and how it produces secular variation. This area of geophysical theory is a most active and challenging one, and it is in rapid flux.

METHODS OF STUDY

The simplest way to detect secular variation is to observe the changing declination of a magnetic compass over some decades; until the twentieth century, that was the only way. All the instruments employed by famous investigators of geomagnetism, from Gilbert in 1600 to Carl Friedrich Gauss in the 1830's, used adaptations of the compass to measure the magnetic parameters and their changes. Among other goals, these scientists aimed to measure these elements more accurately, so as to reveal secular change in shorter time intervals. During the twentieth century, however, there was a sustained trend to replace traditional magnetic needle instruments with ones based on other applications of electromagnetic principles.

Around 1900, research-quality Earth inductors were developed to replace the dip needle and circle in measuring inclination. The idea behind this first of the new electrically based geomagnetic instruments is simple. Rotate a coil of wire about its own diameter in a magnetic field. If the rotational axis differs from the direction of that field, an electric current is induced in the coil, but if the axis coincides with the field, the current will cease. This "null" method now is used to measure inclination more accurately and easily.

The Earth inductor was followed in the 1930's by the flux-gate magnetometer. This instrument is based on a high-permeability alloy, that is, one that magnetizes readily. Around two cores of such material are wound two coils of wire, in opposite directions, that carry the same alternating current, so that the same reversing magnetic field is produced in both cores, but 180° out of phase. When placed in the Earth's field, the changes in the magnetic fields of these two cores do not cancel out, and this changes the current flowing in the coil around each core by different amounts. The net current is related to the component of the Earth's magnetic element in the direction the magnetometer is pointing. When, however, this magnetometer is oriented parallel to the Earth's field, no current is produced. The flux-gate magnetometer has seen wide use in aerial geomagnetic surveys.

Other generations of magnetic instruments have appeared since the flux-gate. Some of the most useful are proton precession magnetometers, rubidium vapor magnetometers, and superconducting magnetometers. These devices take advantage of principles of quantum physics. Some of them, like the proton precession instrument, measure only the total intensity of the Earth's field. Others, like the superconducting magnetometer, are directional. Both types are many times more sensitive than older instruments and also perform much faster.

Magnetic surveys have been an essential part of the method of studying secular variation. All over the world, teams of observers have established "repeat stations," or places for careful observation of the magnetic field parameters at various time intervals. Magnetic surveys have been greatly facilitated not only by the new instruments mentioned above but also by the way those instruments are used. Surveys are now often conducted very quickly with instruments carried by airplanes and satellites (such as MAGSAT). Data that once took decades to gather are now collected in months. Moreover, the extensive calculations needed to analyze global data have been greatly accelerated by computers. Worldwide magnetic charts are produced much more frequently now than in 1900, and the study of secular variation is thus much more detailed.

Equally impressive changes have been wrought by the use of geomagnetic methods to study the magnetic properties of rocks. Igneous and sedimentary rocks that contain iron grains can record the Earth's magnetic field at the time they formed. The phenomenon is called remanent magnetism or paleomagnetism. Until the development in the middle of the twentieth century of techniques for measuring remanent magnetism in rocks, secular variation studies were limited to data obtained by direct measurement of the Earth's field during historic times. Little was known of the magnetic field before 1600. Past phenomena that have been revealed by these methods include reversals of the magnetic field polarity and geomagnetic excursions.

The study of geomagnetism has come a long way with the rapid development of new instruments and methods. No longer is the purpose restricted to just a description of the main field and its variations. With the new sensitivity and portability made possible by electronics, geomagnetic secular variation has become a useful tool in many diverse scientific endeavors, such as archaeological dating of artifacts, magnetostratigraphic dating of sediments, determining rates of seafloor spreading, and tracing continental drift, in addition to the traditional effort to understand processes occurring in the Earth's core.

CONTEXT

Most people are familiar with the magnetic compass, and many know at least roughly how to use it. Two activities which demand close attention to magnetic declination and its secular variation are reading topographic maps and navigating at sea.

In the margin of topographic maps, there usually are arrows which point to true north and magnetic north. With this declination information, one can relate directions on the map to compass readings in the field. In areas where the secular variation of declination occurs rapidly, it is also necessary to know when declination readings were last measured and the rate of their change. For example, near Tay River in the Canadian Yukon, the declination was listed as 33°, 25 arc minutes, east of north in 1979 and decreasing at 3.3 arc minutes per year. Thus, if the secular variation there continued at that rate, in a century the declination would change by 5°, 30 arc minutes, to 27°, 55 arc minutes east of north. Secular variations cannot be predicted reliably over so long a period, however, and maps are therefore updated regularly in magnetic surveys.

Information regarding declination at sea and especially near the coast is of even greater importance. Every ship is sometimes beset by fog, and thus an essential bit of navigational data is the present declination. Up-to-date charts are, again, the best means to avoid dealing with secular variation. As the date of the magnetic declination recedes into the past, however, reliable information concerning its secular change becomes more important.

The deep interior of the Earth is inaccessible to direct study. Thus scientists must watch closely for clues received at the Earth's surface about the conditions and processes in the interior. Magnetic secular variation is one of the ways information can be obtained about the geodynamo. The geomagnetic changes occurring over geologic time have been preserved in some igneous rocks as they cooled and some sedimentary as they settled polarity reversals of the Earth's main magnetic field. This remanent magnetism or paleomagnetism records what the field was like in the past. Such data gathered around the world from rocks of various ages reveal that the geomagnetic field has reversed its polarity many times in the geologic past. These geomagnetic reversals typically occur at intervals of tens of thousands to millions of years; the last one happened about 700,000 years ago. The geomagnetic reversals can be tied in to the chronology of the Earth and are an important element in plate tectonics.

Gregory A. Good

FURTHER READING

Backus, George. *Foundations of Geomagnetism*. New York: Cambridge University Press, 1996. Describes in detail the mathematical and physical foundations of geomagnetism. Technical and more advanced than introductory texts.

De Bremaecker, Jean-Claude. *Geophysics: The Earth's Interior*. New York: John Wiley & Sons, 1985. This well-written text is intended for college-level students with some calculus and some physics background. Nevertheless, the author is careful to explain difficult concepts or mathematical statements. Chapter 10, "Magnetostatics," and chapter 11, "The Earth's Magnetic Field," can be read separately from the rest of the book to provide an in-depth survey of geomagnetism, its measurement, and its secular variation. Especially useful are the technical appendixes on mechanical quantities, magnetic quantities, data about the Earth, notation, and some relevant mathematics. One of the best treatments available.

McConnell, Anita. *Geomagnetic Instruments Before 1900*. London: Harriet Wynter, 1980. This short book provides one of the clearest expositions of the basics of geomagnetism for the lay reader. Includes illustrations of many of the basic early forms of instrumentation, especially European.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. The chapters on the Earth's interior and plate tectonics have sections on the Earth's magnetic field and its reversals in the geologic past.

Thompson, Roy, and Frank Oldfield. *Environmental Magnetism*. London: Allen & Unwin, 1986. This book captures the broad range of possible applications of knowledge of magnetism in the study of the Earth that have appeared since the 1950's, from the study of magnetic minerals to biomagnetism. This is an introductory, nonmathematical, college-level text. Although its chapters on basic magnetic principles are valuable, the most unusual feature of the book is the many application chapters. Especially relevant to secular variation are chapters: "The Earth's Magnetic Field"; "Techniques of Magnetic Measurements"; "Reversal Magnetostratigraphy"; and "Secular Variation Magnetostratigraphy."

See also: Auroras; Earth-Moon Relations; Earth-Sun Relations; Earth's Magnetic Field: Origins; Earth's Magnetic Field at Present; Earth's Magnetosphere; Planetary Magnetospheres; Planetary Orbits: Couplings and Resonances; Van Allen Radiation Belts.

Earth's Magnetic Field at Present

Category: Earth

The study of the Earth's magnetic field is important from both academic and practical perspectives. The Earth is the only planet of the inner solar system with a strong magnetic field, and this provides clues about the formation of the Earth and the other inner planets. The Earth's magnetic field deflects and traps high-energy charged particles, providing a shield to protect life. It can disrupt modern communication and electrical systems, but it also can point to the location of ore deposits.

OVERVIEW

The study of the Earth's magnetic field is a branch of geophysics, which combines geology and physics to investigate various physical characteristics of the Earth. The ultimate source of any magnetic field is moving electrical charge, such as an electric current flowing in a wire. Approximately 90 to 95 percent of the Earth's magnetic field is thought to be produced by electrical currents in the Earth's molten metallic outer core, a mechanism referred to as the geodynamo.

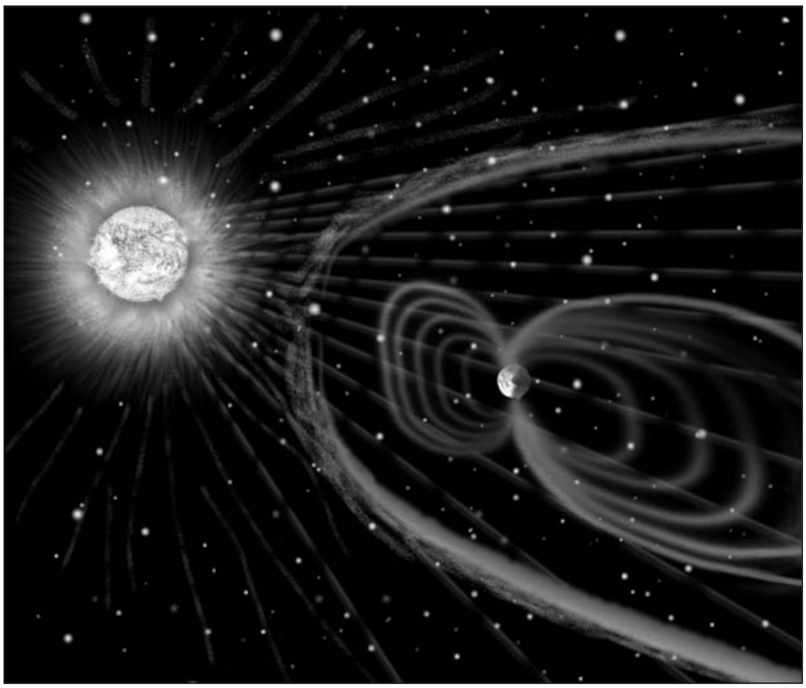
The Earth's field is predominantly a dipole field, meaning it has two magnetic poles; the prefix "di" is derived from the Greek word meaning "two." This is the type of field produced by a bar magnet or an electric current flowing in a wire loop. By definition, the north pole of a bar magnet is the end that points northward on Earth at the present time. Since like magnetic poles repel and unlike magnetic poles attract, that means the Earth's magnetic south pole is located in the Northern Hemisphere, and the Earth's magnetic north pole is located in the Southern Hemisphere. Magnetic field lines are a way of visualizing the direction of a magnetic field, and by convention they point in the direction the north pole of a bar magnet would point. Magnetic field lines leave the Earth's surface in the Southern Hemisphere, arc over the Earth, and reenter the Earth in the Northern Hemisphere. The magnetic poles are the two places where the field lines leave and enter the Earth's surface precisely vertically.

The pole in the Northern Hemisphere is called the north magnetic pole (but remember it is a magnetic south pole), and the pole in the Southern Hemisphere is called the south magnetic pole (although it is a magnetic north pole). At first this may seem confusing, but note the distinction in terminology: the geographic hemisphere that the pole is in comes before the words "Magnetic Pole," while the type of pole comes between words "magnetic" and "pole." The field's strength is about 0.6 gauss (a unit of magnetic induction) at the magnetic poles and about 0.3 gauss at the magnetic equator, where the field lines are horizontal. (For comparison, a small bar magnet has a field strength of about 1 gauss.) The difference in strength is due to the field lines bunching together at the magnetic poles and spreading apart at the magnetic equator.

The magnetic poles are not located at the geographic or rotational poles of the Earth, which are the two points where the rotational axis of the Earth intersects the surface. Currently, near the beginning of the twenty-first century, the north magnetic pole is located in the Arctic Ocean north of Canada and west of Greenland, approximately 900 kilometers from the geographic North Pole. The south magnetic pole is located in the ocean between Antarctica and Australia, approximately 2,900 kilometers from the geographic South Pole. Notice that the magnetic poles are not symmetrically located in relation to the rotational poles.

The magnetic poles also are not stationary; rather, they wander around the polar regions at varying speeds of up to tens of kilometers per year. Since the early 1900's, the north magnetic pole has moved roughly north-northwest about 1,300 kilometers, while the south magnetic pole has moved from Antarctica northward out into the ocean toward Australia. Because the geodynamo, the theoretical source of the Earth's magnetic field, is driven partly by the Earth's rotation, it is presumed that over long periods of time, the positions of the two magnetic poles average out to roughly the locations of the geographic poles. In addition, measurements of the field's strength since the mid-nineteenth century indicate that it is decreasing at a rate of about 6 percent per century. Archaeomagnetic evidence indicates the field was approximately twice as strong two millennia ago, and before that, around 3,500 B.C.E., it was only about one-half the present strength.

These changes in direction and strength over timescales of years to millennia are called secular variations. They are thought to be due to changes in the geodynamo operating in the Earth's molten outer core. Considering its past behavior, scientists cannot predict what the magnetic field will do in the future. It may continue to decrease, or it may increase. If it were to



An artist's conception of the outflowing solar wind meeting Earth's magnetic field. (NASA/ESA) continue decreasing at the present rate, the field would drop to zero in about 1,600 years. This might lead to a magnetic reversal, in which the field re-forms but with its polarity reversed. Paleomagnetic measurements of past magnetic fields preserved in some rocks indicate this has occurred many times in the geologic past, the last time about 700,000 years ago.

The Earth's magnetic field also exhibits small, rapid changes in direction and strength over periods of hours to days, due to a variety of external effects. For example, the gravitational fields of the Sun and Moon distort the atmosphere of the Earth, in the same manner as ocean tides. Movement of electrically charged particles in the atmosphere produces a weak contribution to the magnetic field that changes with the relative positions of the Sun and Moon.

The Sun continually blows electrons, protons, and other electrically charged particles outward from its surface at speeds of hundreds of kilometers per second, a phenomenon known as the solar wind. When these charged particles encounter the Earth's magnetic field, they interact with it, producing a boundary called the magnetopause. Inside the magnetopause is the magnetosphere, the region in which the Earth's magnetic field is dominant. The solar wind changes the shape of the Earth's field. The side facing the Sun is pushed in toward the Earth by the solar wind so that the magnetopause is about 60,000 kilometers, or 10 Earth radii, from the Earth, while the field pointing away from the Sun is elongated into a magnetic tail that can extend farther than the orbit of the Moon.

Some of the solar wind particles, particularly electrons and protons, are trapped by the Earth's magnetic field. These form the Van Allen belts, which were discovered in 1958 by Dr. James Van Allen while analyzing data from a charged particle detector he had placed aboard Explorer 1, the first successful U.S. satellite. The inner belt is a torus about 3,000 kilometers above the magnetic equator; the outer belt is a larger torus about 14,000 kilometers above the magnetic equator.

The number of sunspots increases and decreases over a cycle of eleven years. Sunspots are just one of the more obvious manifestations of solar magnetic activity, and the Sun reverses magnetic polarity with each eleven-year cycle. During times of maximum solar activity, solar flares are most likely to erupt from its surface. These flares eject large numbers of highly energetic, electrically charged particles out into the solar system. If they encounter the Earth's magnetic field, they can produce magnetic storms that cause wild variations in the Earth's field. This in turn can disrupt modern communication and electrical distribution networks. It is at these times, when the Sun is most active, that auroras (the northern and southern lights) are most common. Increased numbers of charged particles from the Sun are deflected by the Earth's magnetic field and enter the Earth's upper atmosphere near the magnetic poles, where they excite air molecules, causing them to glow.

Lightning is a very rapid electrical discharge in the atmosphere; electrical charges can flow from the ground to clouds, from clouds to the ground, or from cloud to cloud. Locally, this strong but brief electrical current produces a very large increase and then decrease in the background field strength.

Magnetic anomalies distort the dipole shape of the main field. Some of these anomalies probably result from more complicated flow patterns in the molten outer core, while others probably are associated with rock units that are rich in iron. Two of the strongest known are located near Kursk, Russia, and in northern Manitoba, Canada. Running parallel to the ocean ridge-rift system are bands or strips of seafloor with alternate normal and reversed magnetic polarity that enhance or weaken the present field over them. The strips preserve a record of the Earth's past magnetic field, frozen into the igneous rocks (mainly basalt and gabbro) that cooled from lava that oozed out along the ridge-rift, and then was pushed away from the ridge-rift as new lava oozed out. This provides evidence of magnetic field reversals in the geologic past and support for the concept of seafloor spreading (one of the key parts of plate tectonics). Small anomalies can even result from man-made iron objects.

METHODS OF STUDY

The orientation of the magnetic field at any point on Earth is specified by two angles called declination and inclination. Declination is the angle between true north (the direction of the geographic or rotational north pole) and the horizontal component of the magnetic field line at that point. Thus declination is the angle between true north and the direction an ordinary compass needle points. Inclination is the angle between a horizontal line and the downward tilt of the magnetic field line at that point. Inclinations are downward (positive) in the northern hemisphere and upward (negative) in the Southern Hemisphere. The magnetic poles are located where the inclination is 90°, specifically 90° down (positive) for the pole in the Northern Hemisphere and 90° up (negative) for the pole in the Southern Hemisphere. The magnetic equator is located where the inclination is 0°.

Around the world, 130 permanent magnetic observatories have been established to record any changes in the magnetic field. It was at observatories in London and Paris that secular variations of the field were first recognized in the 1600's. Early observatories could measure only the declination and inclination of the field. Declination was measured with a compasslike device and inclination with a magnetized rod balanced so that it could pivot freely in a vertical plane.

Magnetometers for the measurement of magnetic field intensity were first developed in the mid-1800's, and a number of different types are in use today. In conjunction with the magnetic observatories on the ground, some satellites carry magnetometers for the measurement of the field from orbit, and they provide readings for virtually the entire globe.

Portable magnetometers can detect local field anomalies due to things under the surface. Geologists use them to prospect for magnetic iron ore deposits, and archaeologists use them to search for buried iron artifacts.

CONTEXT

When magnetic storms occur, modern communication and electrical distribution networks can be disrupted. Also on these occasions, auroras are more likely to occur and be seen over larger areas. The magnetic field interacts with electrically charged particles and prevents many of them from reaching the Earth's surface. It is possible that a decrease in the field would lead to more particles reaching the surface, perhaps producing greater numbers of genetic mutations or cancers. Changes in the field strength have been suggested as a cause of some of the mass extinctions that have occurred in the geologic past.

Stephen J. Shulik

FURTHER READING

Courtillot, V., and J. L. Le Mouel. "Time Variations of the Earth's Magnetic Field: From Daily to Secular." In *Annual Review of Earth and Planetary Sciences* 16 (May, 1988): 389-486. This source covers the geomagnetic field in general and then provides an in-depth study of its variations, from short-term to very long-term. Very little mathematics; many figures.

Fowler, C. M. R. *The Solid Earth: An Introduction to Global Geophysics*. 2d ed. New York: Cambridge University Press, 2004. An updated version of a widely used textbook for introductory geophysics courses. Designed for students with some knowledge of physics and calculus.

Garland, G. D. *Introduction to Geophysics*. 2d ed. London: W. B. Saunders, 1979. Used as a text for introductory geophysics, this book contains in chapter 17, "The Main Field," readable material on the main field and its generation. Time variations are discussed as well as the external field and methods of measurement. Some equations; many figures and graphs of interest to the less specialized reader.

Jacobs, J. A., R. D. Russell, and J. T. Wilson. *Physics and Geology*. 2d ed. New York: McGraw-Hill, 1974. This introductory geophysics textbook is formidable for the average student because it uses considerable mathematics in some chapters, but chapter 8, "Geomagnetism," has sections on the present field and contains a minimum of equations and many figures and graphs. Auroras and the magnetosphere are also discussed.

Knecht, D. J., and B. M. Shuman. "The Geomagnetic Field." In *Handbook of Geophysics and Space Environment*, edited by A. S. Jursa. Springfield, Va.: National Technical Information Service, 1985. This source covers the geomagnetic field and various aspects of it: terminology, sources of the field, measurements, the main field, and sources of geomagnetic data. Some sections have no mathematics, but others have a small amount. Many figures help the reader to understand the authors' narratives. A number of references are listed at the end of the chapter.

Motz, Lloyd, ed. *Rediscovery of the Earth*. New York: Van Nostrand Reinhold, 1979. As a collection of articles for the nonscientist by scientists renowned in their respective fields, the text makes interesting reading, augmented with many colorful illustrations. The chapter "The Earth's Magnetic Field and Its Variations" is written by Dr. Takesi Nagata, who has written hundreds of articles on diverse aspects of geophysics besides the Earth's magnetic field.

Smith, David G., ed. *The Cambridge Encyclopedia of Earth Sciences*. New York: Cambridge University Press, 1982. Chapter 7, "The Earth as a Magnet," contains information about the Earth's present-day magnetic field, geomagnetic field changes, and magnetic anomalies. The text is well written at a nontechnical level, with many colorful diagrams and figures.

Stacey, F. D. *Physics of the Earth*. New York: John Wiley & Sons, 1977. In section 8.1, "The Main Field," the author provides a short, technical description of the main field that is of interest to the more advanced student. As a textbook for geophysics, it covers many other areas on a technical level.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. The chapter on the Earth's interior has a section on the Earth's magnetic field.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996. Covers geophysics research and theories since 1960. Includes information about Pangaea, the supercontinent cycle, and the reversals of the Earth's magnetic field. For general audiences.

See also: Auroras; Earth-Moon Relations; Earth-Sun Relations; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetosphere; Planetary Magnetospheres; Planetary Orbits: Couplings and Resonances; Van Allen Radiation Belts.

Earth's Magnetosphere

Category: Earth

The Earth's magnetosphere is the region around the Earth in which the geomagnetic field is stronger than the interplanetary magnetic field. The outer boundary of the magnetosphere is termed the magnetopause. The flow of charged particles from the Sun, called the solar wind, pushes in the magnetopause on the side of the Earth facing the Sun in a bow shock effect and drags it out into a long magnetotail on the side facing away from the Sun. The Earth's magnetic field interacts with the charged particles of cosmic rays as well as those emanating from the Sun, protecting life from their harmful effects.

OVERVIEW

The magnetic field at Earth's surface has been used in navigation for centuries. However, studying the extent and behavior of the geomagnetic field out in space around Earth became possible only with the dawn of the space age. The launch on January 31, 1958, of Explorer 1, the first successful U.S. satellite, brought a completely unexpected discovery. Belts of electrically charged particles circling Earth and trapped by Earth's magnetic field were detected by James Van Allen when he analyzed the signals sent back from a Geiger counter he had placed on board the satellite. These belts were named the Van Allen radiation belts in his honor. Since Explorer 1, myriad spacecraft have explored Earth's magnetosphere, measuring the field's strength and direction, the numbers and energies of the charged particles in it, and its interaction with the solar wind and interplanetary magnetic field.

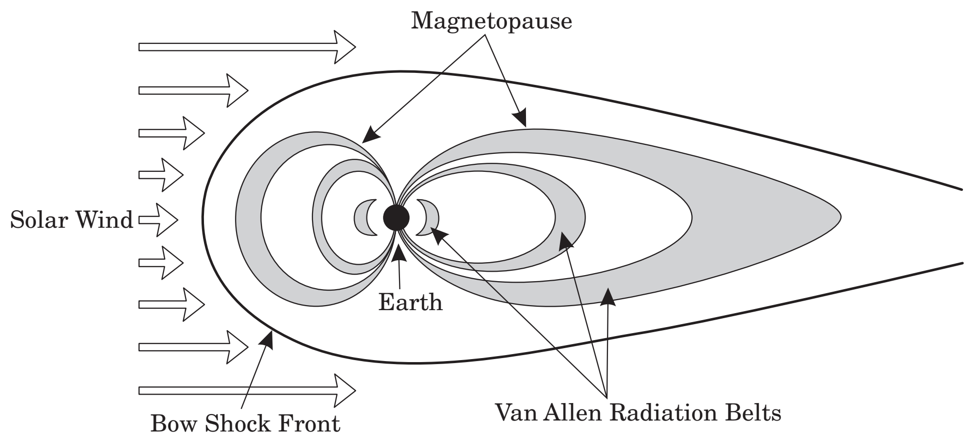
The magnetic field near Earth is basically a dipole field, the type of field produced by a bar magnet or an electric current flowing in a wire loop. It is as if a bar magnet were inside Earth, tilted about 10° to 15° with respect to the axis of rotation and offset from Earth's center several hundred kilometers toward the Pacific Ocean. The north pole of this hypothetical magnet actually is located in the Southern Hemisphere and is called the south magnetic pole; the south pole of the magnet is located in the Northern Hemisphere and is called the north magnetic pole. The north magnetic pole (actually a magnetic south pole) is located north of Canada and west of Greenland about 8° away from the geographic or rotational north pole, and the south magnetic pole (actually a magnetic north pole) is located off the coast of Antarctica in the direction of Australia about 26° from the geographic or rotational south pole. This somewhat confusing situation and terminology is because opposite poles attract, and by definition a magnet's north and south poles are the north-seeking and south-seeking ends of the magnet used as a compass.

Although imagining a bar magnet inside Earth is an easy way to visualize the geomagnetic field, this cannot be the actual situation, since Earth's interior is much too hot for any material to be permanently magnetic. Instead, the source of the geomagnetic field and magnetosphere is thought to be fluid motions and electric currents in Earth's molten metallic outer core, in what is termed the geodynamo process.

With increasing distance from Earth, the geomagnetic field lines become distorted from those of a simple bar magnet by the solar wind—the flow of electrically charged particles blown from the Sun's surface out into interplanetary space. The flow of charged particles carries the solar magnetic field, which becomes the interplanetary magnetic field. The solar wind pushes back the geomagnetic field lines on the side of Earth facing the Sun and stretches them out into a magnetotail of nearly parallel field lines on the side away from the Sun. The boundary where the geomagnetic field equals the solar/interplanetary magnetic field is called the magnetopause; inside it is Earth's magnetosphere, where the geomagnetic field dominates. The magnetosphere typically extends out to about 65,000 kilometers on the side of Earth facing the Sun and up to millions of kilometers in the magnetotail pointing away from the Sun.

Where the rapidly moving charged particles of the solar wind encounter the magnetosphere, their speed abruptly decreases and their density abruptly increases, forming a shock wave called a bow shock, analogous to the water-wave pattern around the bow of a rapidly moving boat. Most of the particles in the solar wind flow

Earth's Magnetosphere



past Earth around this bow shock and drag the magnetosphere out into its long magnetotail. The magnetopause and its accompanying bow shock fit around Earth like a sock with a golf ball in its toe.

Some of Earth’s field lines originating near a magnetic pole extend through the magnetopause into the magnetosheath, the region between the magnetopause and the bow shock. These field lines form the polar cusps or clefts where they pass through the magnetopause. It is along these field lines that some of the charged particles in the solar wind can enter Earth’s magnetosphere through the polar cusps.

The magnetic field lines that exit the surface in one hemisphere and bend around Earth and reenter the surface in the other hemisphere are called closed field lines. Those that do not return but extend through a polar cusp into the magnetosheath or continue out into the magnetotail are called open field lines. At the magnetic poles, the boundary between the open field lines and the closed field lines is called the auroral oval. Because the auroral oval is the dividing line between closed field lines where particles may be trapped and open field lines where trapping is impossible, it is where the energetic charged particles of the solar wind can enter Earth’s atmosphere and produce auroras. Spacecraft images have shown that the auroras are fairly continuous phenomena around the auroral ovals. Charged particles also can escape through the auroral ovals; in particular, streams consisting mainly of helium ions—the polar wind—flow out to distant regions of the magnetosphere.

Along the central plane of the magnetotail runs a flattened plasma sheet of ions (atoms that have lost one or more electrons) and electrons, separating the magnetic field lines originating from Earth’s north and south magnetic poles. Currents flow crosswise in the plasma sheet and maintain the magnetic field separation. At great distances from Earth, the magnetotail may sometimes undergo a magnetic field line reconnection or collapse. This releases a large amount of stored magnetic field energy, which in turn accelerates the charged particles in the magnetosphere. Such an event is called a geomagnetic substorm. During a substorm, high-energy electrons enter the atmosphere near the auroral oval, producing particularly strong auroras and disrupting the ionosphere and long-range radio communications, which are reflected by the ionosphere. At these times, ring currents flow around the magnetic equator at a distance from Earth of about 12,500 kilometers. Magnetometers on the ground can measure changes in the magnetosphere produced by these currents.

At times of increased solar activity, near the maximum of the solar sunspot cycle, solar flares send out strong bursts of charged particles in the solar wind, which can strongly compress the magnetosphere on the Sun side of Earth and collapse the magnetotail on the far side. Then the magnetopause may shrink to within 12,000 kilometers of Earth. When that occurs, geosynchronous satellites orbiting at a distance of about 40,000 kilometers (where the satellite revolution period equals Earth’s rotation period) enter the magnetosheath and become subject to interactions with solar wind particles, causing the satellites to charge suddenly to high voltages. Some communications satellites have been disabled or have had their signals disrupted by internal sparks at such times.

When charged particles such as electrons or ions move in a magnetic field, they follow corkscrew paths around the magnetic field lines. As a charged particle corkscrews around a magnetic field line, the angle its path makes with the field line (its “pitch angle”) increases as it moves into regions of greater magnetic field strengths, until it is moving in a circle perpendicular to the field line and can go no farther in the direction of the field. It has reached its “mirror point” and must now return in the direction from which it came.

The field lines of a magnetic dipole reach out from regions of high field strength near one pole to regions of weaker intensity at the magnetic equator, and then return to regions of high field strength near the opposite pole. Consequently, a particle corkscrewing around one of these field lines will be mirrored near one magnetic pole, travel back to the vicinity of the other pole and be mirrored again, and so on. High-energy particles can reach farther in toward the pole than low-energy particles, and they may even enter the atmosphere, producing the beautiful auroras, or northern and southern lights. Less energetic particles are trapped on the magnetic field lines outside the atmosphere, and they can only escape by colliding with other particles or by slowly moving across field lines in response to electric fields or changing magnetic fields. As Earth rotates, these trapped particles swing around with the rotating magnetic field.

KNOWLEDGE GAINED

Earth's magnetosphere is made up of Earth's magnetic field and the charged particles controlled by it. The magnetosphere and magnetic field can change over hours, days, or years in response to currents of charged particles and the surrounding interplanetary magnetic field.

Earth has two invisible belts of trapped energetic particles, the Van Allen belts. They were discovered by James Van Allen in 1958 using data from the first successful U.S. satellite, Explorer 1. The inner belt, which extends from 1,000 to 5,000 kilometers above the equator, is kidney-shaped in cross section and contains mainly high-energy protons. The outer belt, 15,000 to 25,000 kilometers from Earth, is crescent-shaped in cross section and contains mainly high-energy electrons. Because the charged particles are trapped in the magnetic field, they cannot easily leave the belts. They pose a danger to astronauts or sensitive electronic equipment orbiting for long periods within the Van Allen belts.

CONTEXT

It is important to understand the processes that occur in Earth's magnetosphere because they strongly affect conditions on Earth. Only since the space age began have the structure and processes of Earth's magnetosphere been subject to study.

Earth's magnetosphere performs the vital function of shielding living things from the possibly harmful effects of high-energy charged particles, such as cosmic rays and those that come from the Sun. When space travelers leave the protection of the magnetosphere, they must be shielded against charged particles, especially the intense bursts of them emitted in solar flares. Furthermore, strong solar flares may disrupt the magnetosphere, damaging or disabling Earth-orbiting satellites, long-distance radio communications, and electricity distribution networks.

At times in Earth's history when its magnetic field became very weak and reversed direction, living things may have been subjected to an intense flux of high energy charged particles, from cosmic rays and solar flare ejections, which could have led to increased rates of cancer and genetic mutation. Genetic mutations can be both good and bad. Although most mutations are deleterious (and even deadly), some are beneficial and convey an evolutionary advantage. Therefore, those times when the geomagnetic field was weak or absent may have contributed to mass extinctions or rapid evolution.

Some theorists propose that changes in the number of charged particles entering Earth's atmosphere, or changes in the solar/interplanetary magnetic field as it interacts with Earth's magnetic field, may influence the weather on Earth. Perhaps, they suggest, auroral currents modify polar air currents, leading to periods of drought or increased precipitation in different regions on Earth.

Dale C. Ferguson

FURTHER READING

Akasofu, S. I., ed. *Dynamics of the Magnetosphere*. Dordrecht, Netherlands: D. Reidel, 1980. A collection of contributions to a 1979 meeting of magnetospheric scientists in Los Alamos, New Mexico. Includes results of experiments and readable, condensed, technical summaries. This volume is mostly concerned with the disturbed magnetosphere, as during geomagnetic substorms. Well illustrated.

Akasofu, S. I., and Y. Kamide, eds. *The Solar Wind and the Earth*. Dordrecht, Netherlands: D. Reidel, 1987. Written by experts, this is a collection of chapters about the Sun and Earth and their interactions. Includes sections on Earth's ionosphere and thermosphere. The book is somewhat technical, but nevertheless clear; the history of each subtopic is well treated. Contains lists for further reading.

Allen, Oliver E. *Atmosphere*. Alexandria, Va.:

Time-Life Books, 1983. A popular book that covers all aspects of the atmosphere. A good treatment of the history of atmospheric studies. The interested layperson will find the relationship of the magnetosphere to Earth's atmosphere well explained. Contains photographs, illustrations, and a bibliography.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Has a section on Earth's magnetosphere.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Has a section on Earth's magnetosphere.

Friedman, Herbert. *Sun and Earth*. San Francisco: W. H. Freeman, 1986. A volume that lucidly describes the Sun's effects on Earth's magnetosphere and ionosphere. Written for the layperson by a pioneer in spacecraft exploration. Of moderate length, the book contains photographs, drawings, and an appendix of references to specific topics.

Hargreaves, John K. *The Upper Atmosphere and Solar-Terrestrial Relations*. New York: Van Nostrand Reinhold, 1979. This textbook delves into the physics of the magnetosphere. It may profitably be read by specialists in the field, other physical scientists, or mathematically inclined students. Includes useful line drawings, numerous equations and references, and questions for study.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college textbook beyond the introductory level, its approach is based on comparative planetology. Provides a succinct but easily understood description of planetary magnetic fields and magnetospheres.

Johnson, Francis S., ed. *Satellite Environment Handbook*. 2d ed. Stanford, Calif.: Stanford University Press, 1965. An excellent technical reference that has lost little value with age. It covers the near-Earth environment, from the magnetic field to micrometeoroids, and has a good section on the magnetosphere. Includes graphs, tables, line drawings, and references. Written by many knowledgeable contributors.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996. Covers geophysics research and theories since 1960. Includes information about Pangaea, the supercontinent cycle, and the reversals of the Earth's magnetic field. For general audiences.

See also: Earth-Moon Relations; Earth-Sun Relations; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Planetary Magnetospheres; Planetary Orbits: Couplings and Resonances; Van Allen Radiation Belts.

Earth's Mantle

Category: Earth

The mantle is the part of the Earth's interior that lies between the crust and the outer core. It is composed of rocks that are of greater density than those of the crust. The mantle generally is solid, except for pockets of molten magma, but it contains a layer called the asthenosphere where the actual temperature is so close to the melting point that the rock there is in a plastic state. It is upon this layer that the Earth's lithospheric plates ride.

OVERVIEW

The mantle is the portion of the interior of the Earth that extends from the base of the crust to the boundary with the outer core. This distance is approximately 2,900 kilometers, roughly 45 percent of the radius of the Earth. Since the thickness of the Earth's crust is not uniform, the distance from the ground surface to the upper boundary of the mantle, called the Mohorovičić discontinuity or simply Moho for short, varies significantly. The thickness of the crust in continental areas averages approximately 40 kilometers and reaches as much as 70 kilometers under young high mountain ranges, while in the ocean basins the crust is only about 5 to 7 kilometers thick.

Evidence for layers of different density within the Earth was first found in 1909 by the Croatian seismologist Andrija Mohorovičić, using seismic waves from earthquakes. Studying an earthquake that had occurred in Yugoslavia, he discovered that seismographs located a few hundred kilometers from the epicenter of the quake had recorded two sets of seismic waves. He correctly deduced that one set had traveled directly from the quake, while the other set had been refracted at a boundary somewhere below the surface where there was an abrupt change in seismic-wave speed. At stations nearer the quake, the direct set of waves arrived first, but at more distant stations, the refracted set of waves arrived first. This meant the speeds below the boundary were faster than those above; from the arrival times of the refracted waves, Mohorovičić was able to determine their speeds below the boundary. From the speeds above and below the boundary and the arrival times of both sets of waves, he could calculate the distance from the surface to the boundary, which has become known as the Mohorovičić discontinuity (or just Moho for short) in his honor. This discontinuity is taken to be the boundary between the crust above and the mantle below.

During the mid-1960's, a project to drill down through the Earth's crust to sample the mantle below the Moho was begun. The undertaking was appropriately named Project Mohole. Because of technical difficulties and cost overruns, the project was abandoned. However, we know of locations where mantle rocks have been brought to the surface by various geologic processes, so, in effect, we have been provided with free samples. The rock peridotite, composed mostly of the silicate minerals olivine and pyroxene, seems to represent the general composition of the mantle. Another likely mantle rock is eclogite, which might transform into basalt in the crust as pressure drops with reduced depth. A third type of mantle rock is kimberlite, which occurs in pipe-shaped deposits and is mined extensively for diamonds. Diamonds are a high-pressure form of carbon that originate in the mantle at depths between 100 and 200 kilometers and then are carried to the surface in kimberlite pipes.

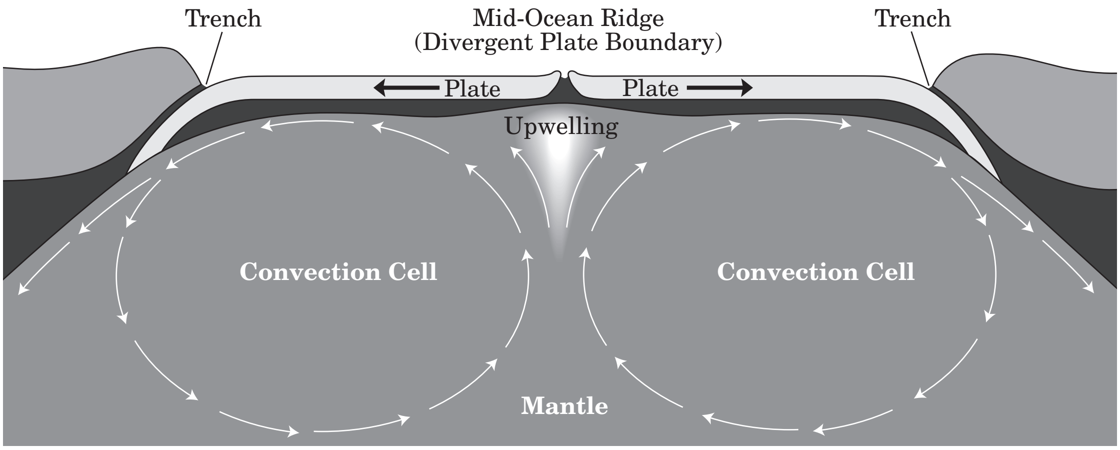
The greatest source of information on the physical nature of the mantle comes from the study of seismic waves. Seismic wave speeds gradually increase through the upper mantle down to a certain depth and then decrease a bit in what is called the low-velocity zone, first identified by Beno Gutenberg in 1926. This low-velocity zone begins about 100 kilometers below continents, but as little as 20 kilometers below ocean ridge-rifts. It is thought this decrease in wave speed is due to a decrease in the rigidity of the rock. From the top of the low-velocity zone down to depths of about 400 kilometers, the rock is likely to be close to (but not quite at) its melting point, so the rock exhibits plastic behavior. This means it is solid, but deformable, and it is capable of flowing or oozing very slowly. This plastic zone is called the asthenosphere.

Above the asthenosphere is the rigid lithosphere, made up of the crust and very top part of the mantle. According to plate tectonic theory, the lithosphere is divided into a number of blocks called tectonic plates. These plates float on the underlying plastic asthenosphere.

Beneath the asthenosphere, wave speed begins to increase again. At depths of about 410 and 660 kilometers, sharp increases in seismic wave speed occur, representing abrupt increases in density. The pressure at 410 kilometers depth is sufficient to collapse the crystal structure of the mineral olivine into the more compact and denser mineral spinel. The even greater pressure at 660 kilometers depth converts the crystal structures of both spinel and pyroxene into the even denser mineral perovskite. Below 660 kilometers, the wave speeds gradually increase down to the boundary with the outer core. (Below this boundary, down in the outer core, primary, or P, waves slow down considerably, and secondary, or S, waves disappear, indicating the outer core is molten.) As a result of the increase in pressure with depth, the density increases from about 3.3 grams per cubic centimeter at the top of the mantle to about 5.6 grams per cubic centimeter at its base.

A critical component of modern plate tectonic theory is the heat flow that drives the motion of the plates. The Earth's interior was heated during its formation about 4.5 billion years ago, but it remains hot today mainly due to the decay of radioactive elements. It is thought that this

Hess's Theory of Convection



heat causes rock in the mantle to slowly flow, rising where it is hotter and less dense, and sinking where it is cooler and denser. It is not known whether this is due to an organized system of convection cells or to an irregular pattern of rising mantle plumes and descending lithospheric plates. The lithospheric plates are carried on the asthenosphere as it flows laterally from areas of upward movement to areas of downward movement.

It has been known since the nineteenth century that the age of the Hawaiian islands increases from southeast to northwest. In the 1960's, when plate tectonic theory was being developed, it was postulated that these volcanic islands recorded the movement of the seafloor over a hot spot or rising mantle plume. Lava from the hot spot is extruded out onto the seafloor, eventually building up to break the surface of the ocean and form a volcanic island. As the Pacific plate moves slowly toward the northwest, the newly created island is carried away from the hot spot, its volcano becomes extinct, and another volcanic island starts to form, creating the Hawaiian island chain. Another such hot spot is located under Yellowstone National Park and fuels all of its geothermal activity.

METHODS OF STUDY

Although the Earth's mantle cannot be directly observed, samples from it have been brought to the surface by geologic processes, and they can be studied in the laboratory to determine chemical composition and physical properties. However, the primary method for studying the mantle is by analyzing seismic waves generated by an earthquake or an explosion. Some seismic waves travel along the surface of the Earth. They are called L waves (L standing for "long" because of their long wavelengths); they do not travel far, but they cause most of the destructive shaking in an earthquake. Other seismic waves are called body waves because they travel through the body of the Earth, and these are the ones that are used as probes of the Earth's interior. There are two types of body waves: P (for primary) waves travel faster and arrive at seismic stations first, while S (for secondary) waves travel more slowly and arrive second. P waves are longitudinal compression waves, similar to sound waves in that they oscillate in the direction of propagation, alternately compressing and extending the material through which they travel. Like sound waves, P waves can travel through any kind of material—solid, liquid, or gas. S waves are transverse shear waves that oscillate from side to side, perpendicular to the direction of propagation. They will travel only through rigid materials, meaning they will travel only through solids, not liquids or gases.

S waves travel through the entire mantle (except for isolated pockets of molten rock called magma chambers), and this indicates the entire mantle is solid (except for the magma chambers). The speeds of both P and S waves generally increase with depth throughout the mantle, due to the increase in density with depth. The exception is in the low-velocity zone that begins at a depth of between about 20 and 100 kilometers. There both P and S waves slow down, not because of a change in density or composition, but rather because of a change in rigidity. In the low-velocity zone or asthenosphere, the temperature is close to the melting point, so the rocks, although solid and not molten, are soft and deformable, and capable of slow flow at rates of up to a few centimeters per year.

Seismic tomography uses seismic waves to construct images of the Earth's interior the same way computerized axial tomography (CAT) scans use X rays to construct images of internal organs for medical diagnoses. A CAT scan is a composite image of X rays of the body taken from a number of different directions, computer-assembled to show slices through the body; these slices can be stacked to produce a three-dimensional view of the internal organs. In a similar fashion, seismic stations all around the world provide data on waves that have traveled through the Earth in many different directions. When such seismic records are computer analyzed and combined, slices through the Earth and three-dimensional representations of the Earth's interior structure can be produced.

CONTEXT

Processes occurring in the Earth's mantle affect the Earth's surface. The role of the mantle is essential in understanding plate tectonics and the phenomena associated with it, such as earthquakes, volcanic activity, seafloor spreading, the movement of continents, and the location of mineral resources. The tectonic plates of the rigid lithosphere ride on the plastic rock of the asthenosphere in the mantle. Most earthquakes and volcanism occur along the boundaries between the plates, where the plates diverge, converge, or move sideways past each other.

For example, Southern California is prone to earthquakes because it is located along the boundary between the North American plate and the Pacific plate; as the Pacific plate slides to the northwest past the North American plate at the rate of several centimeters per year, sudden slipping of the rocks along the boundary produces earthquakes. At convergent boundaries where one plate slides back down into the mantle in a process called subduction, earthquakes and volcanism are common. The subducting plate rubs against the other plate as it descends under it, producing earthquakes. As it descends to greater depths with hotter temperatures, the minerals in it with lower melting temperatures differentially melt and rise as molten blobs of magma; if they erupt at the surface, they fuel volcanic activity. The volcanoes and earthquakes of Japan, the Philippines, the Andes Mountains along the west coast of South America, and the Cascade Mountains of the northwestern United States all are due to the subduction of one plate under another.

The mechanism that causes the movement of the lithosphere upon the asthenosphere is heat. Hot rock slowly rises from deep in the mantle, gives up its heat while moving laterally, and then sinks back down. This "conveyor belt" motion in the mantle moves the lithospheric plates above.

Processes occurring along plate boundaries as well as at isolated hot spots or mantle plumes can concentrate mineral resources into economically important deposits. Understanding these processes and being able to recognize where they occurred in the past are important in locating such valuable deposits.

David W. Maguire

FURTHER READING

Fowler, C. M. R. *The Solid Earth: An Introduction to Global Geophysics*. 2d ed. New York: Cambridge University Press, 2004. An updated version of a widely used textbook for introductory geophysics courses. Designed for students with some knowledge of physics and calculus.

Jacobs, John A., Richard D. Russell, and J. T. Wilson. *Physics and Geology*. 2d ed. New York: McGraw-Hill, 1974. A technical volume covering such topics as composition of the Earth, geochronology, isotope geology, thermal history of the Earth, magnetism, and seismic studies. The text is intended for college-level students of geology or physics. Some differential equations are used in the book.

Skinner, B. J., and S. C. Porter. *The Dynamic Earth*. 5th ed. New York: John Wiley & Sons, 2006. A well-written, well-illustrated, colorful volume on the geology of the Earth. It would be suitable for the college student beginning geology.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. There is a good chapter on the Earth's interior, as well as several chapters on volcanoes, earthquakes, and plate tectonics that describe their relationships with the structure and processes of the mantle.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996. Covers geophysics research and theories since 1960. Includes information about Pangaea, the supercontinent cycle, and the reversals of the Earth's magnetic field. For general audiences.

Weiner, Jonathan. *Planet Earth*. New York: Bantam Books, 1986. A colorful, well-illustrated, well-written book describing the Earth and how it is studied. This volume is the companion to the PBS television series of the same name. It is suitable for general readers.

See also: Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Oceans; Earth's Origin; Earth's Structure; Planetary Tectonics.

Earth's Oceans

Category: Earth

Earth's ocean water was derived by outgassing from hydrated minerals bound up during the formation of the Earth. Subsequent evolution of the water primarily involved ions from continental and oceanic bottom sediments dissolving in the fluid medium to yield the basic saltiness characteristic of Earth's oceans.

OVERVIEW

Of all the planets in the solar system, Earth stands out as basically a watery world, distinguished from the other planets by large quantities of liquid water. In all, Earth has about 1.36 billion cubic kilometers of water, and 97.2 percent is stored in the oceans. The remaining 2.8 percent of Earth's water not in the oceans is apportioned among ice (77 percent of the total remaining water) and continental and atmospheric waters. The ice itself, now principally in the Arctic-Greenland area (1.72 million square kilometers, up to 3,200 meters thick) and the Antarctic area (12 million square kilometers, up to 4,000 meters thick), has effects ranging from climate control to providing habitats for living organisms to being a reservoir for water that, when added to or removed from the oceans in the past, has caused sea level to rise or fall more than 100 meters.

Ocean water is salty because it contains dissolved minerals; with a salinity of 35,000 parts per million, there is enough dissolved salt to cover the entire surface of the Earth to a depth of about 50 meters. This salty solution is composed primarily of sodium and chlorine ions (together constituting about 86 percent of the ions by weight), along with ions of magnesium, calcium, potassium, and sulfate and carbonate groups. Seawater is slightly basic, with a pH of 8 for the hydrogen-ion concentration.

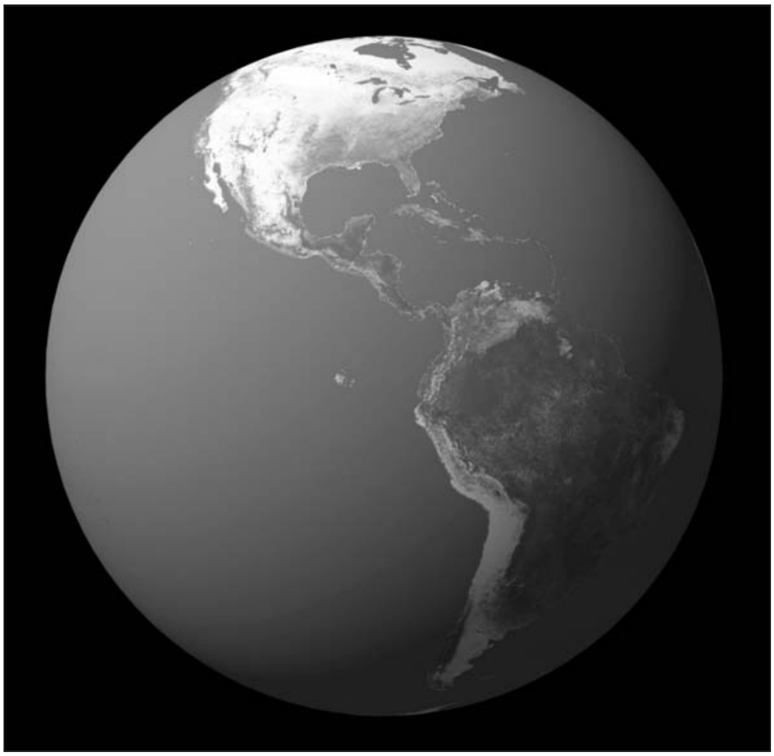
The problem of the oceans' origin is twofold: (1) the primordial origin of the water itself and (2) the origin and rate of addition of the ions that make the oceans salty. The database for solving these problems includes the chemistry of water, the amounts and types of runoff delivered by rivers into the sea, and the composition of volcanic gases, geysers, and other vents opening to the surface, since the oceans and atmosphere are linked in origin.

Numerous sources for the Earth's water have been proposed, and the problem has not been resolved completely. Possible sources include the primordial solar nebula, the solar wind acting over time, delivery or impact degassing by bodies colliding with the Earth, and outgassing from the planetary interior. Processes controlling water on Earth involve rates, amounts, and types of outgassing, modes of planetary formation, possible chemical reactions providing water, loss rates of gases to space, and, finally, internal feedback mechanisms such as changes in Earth's albedo (reflectivity), temperature, alteration of materials, and other factors not clearly understood.

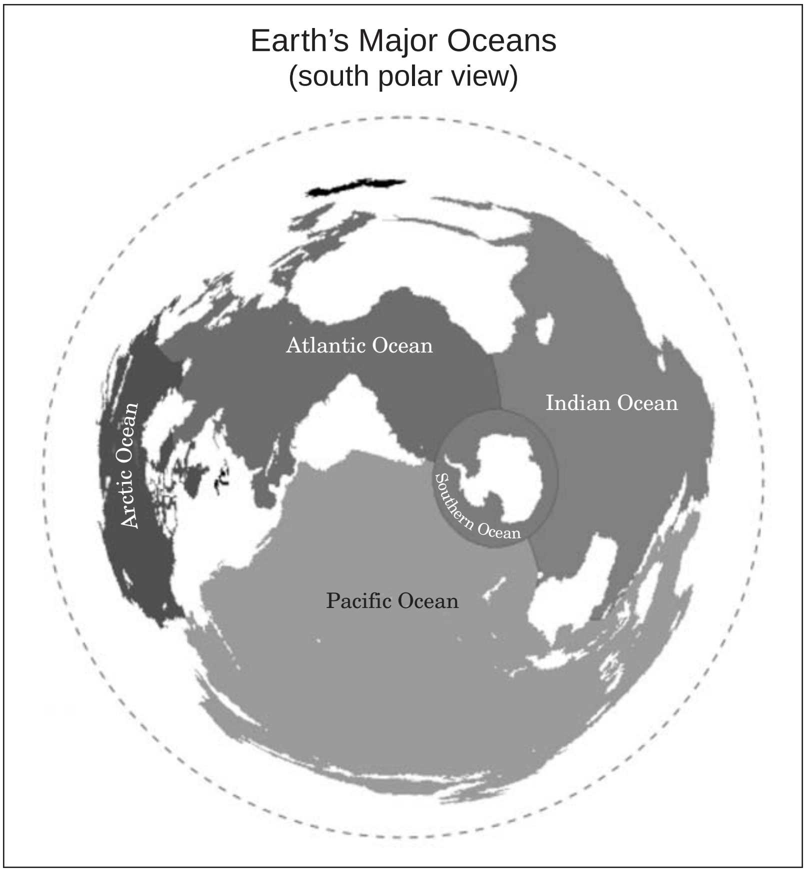
The solar wind as a primary source can be eliminated for several reasons. The basic constituents, protons, may help form water in the atmosphere by reactions with oxygen, but all evidence points to no free oxygen in the primordial atmosphere. The geologic record shows the presence of liquid water at least 4 billion years ago, but Earth was substantially devoid of free oxygen then.

Colliding bodies would be from two primary sources: meteorites and comets. The basic composition of cometary nuclei, consisting of water ice, various ions, metals, organic molecules, and dust grains, would supply enough water, provided that gigantic numbers of cometary objects struck the Earth during the first half-billion years of history. No conclusive evidence for such happenings is available at present, although a theory that Earth is still being bombarded incessantly by small comets containing large quantities of water is supported by the detection of diffuse ice balls entering the atmosphere in the 1990's. Meteoritic impact, particularly during the early stages after final planetary accretion, would definitely add water to the crust via two mechanisms. Carbonaceous chondrites, the oldest and most primitive meteorites, contain abundant volatiles, such as water, chemically bound in various minerals. Additional water, trapped in crustal and mantle rocks since Earth's accretion and differentiation, would have been released during impacts, especially by large impacting objects. It has been calculated that such impact degassing could have released 10^{22} kilograms of volatiles, quite close to the currently estimated value of 4×10^{21} kilograms for the Earth as a whole. Remnants of such ancient astroblemes are lacking, however, because of subsequent erosion, filling in by molten lava, or shifting of the continental masses over 4 billion years.

The most widely accepted origin for the



This image of Earth's Western Hemisphere, compiled from images gathered by NASA's SeaWiFS instrument, shows our planet as a world dominated by water. (NASA/GSFC/SeaWiFS Project/GeoEye, Scientific Visualization Studio)



oceans and atmosphere combines the features of the primordial solar nebula and slow outgassing from within the solidifying Earth. Original water would have been combined with silicate minerals and metallic materials during the planetary accretion process, the hydration assisted by the heating of the Earth due to infalling bodies and the decay of radioactive elements. Such wet silicates appear to be able to hold large quantities of bound water for indefinitely long periods of time. The primordial Earth, believed to have accreted cold, trapped water molecules. If Earth had started too hot, all the minerals would have been dehydrated, and if too cold, no water would have been released; a delicate balance of temperature must have been achieved. Further, the volatiles forming the atmosphere must have outgassed first, in order to provide an insulating blanket under which water could form a liquid phase.

A secondary problem deals with how swiftly the fluids outgassed, either all at once, as many individual events, or in a continuous fashion. Most studies suggest the continuous mode of emission, with greatest reliance on data from currently active sources, such as volcanoes, undersea vents, and associated features. Fumaroles, at temperatures of 500-600° Celsius (800-900 kelvins), emit copious quantities of water, sulfur gases, and other molecules. These structures grade gradually into geysers and hot spots, areas where water is moved crustward from great depths. Magmatic melts rising in volcanoes release water and other gases directly to the surface. In Hawaii, for example, the Halemaumau Pit, the most active vent on the volcano Kilauea, emits 68 percent water vapor, 13 percent carbon dioxide, and 8 percent nitrogen, with the rest mostly sulfurous gases. Similar values are found for ocean-ridge-axis black and white smokers, where hydrothermal accretions result in spectacular deposits of minerals falling out of solution from the emerging hot mantle waters. Detailed studies show that water is trapped in the altered minerals within the basaltic crust of the oceanic plates, 5 percent of the rocks (by weight) in the upper 2 to 3 kilometers being water and hydroxide ions. Free water is known to be extremely buoyant, rising in the crust along shallow dipping faults. Bound water, subducted to great depths, would be expected to cook, moving upward as the rock density lessens and then acting as a further catalyst for melting the surrounding rocks.

In the Earth's earliest stage, the primordial atmosphere was released, only to vanish from the Earth because of overheating. In the second stage, gases were released from molten rocks, with a surface temperature of 300° Celsius (600 kelvins), providing 70 percent water and large quantities of carbon dioxide and nitrogen. In stage three, the atmosphere and oceans gradually changed as a result of volcanoes and weathering action; more and more water was deposited as liquid as the temperature fell. Then the atmosphere added oxygen either by thermal dissociation of water molecules, photochemical breakdown of high-altitude water, or photosynthetic conversion of carbon dioxide to oxygen in plants.

The saltiness of the oceans can be accounted for by the extreme dielectric constant of water, essentially ensuring that ocean water does not remain chemically pure. Geologic evidence shows the general composition of seawater to be similar over time, the content stability attributable to the continuous seawater-sediment interface. John Verhoogen has shown that only 0.7 percent of the ocean has been added since the Paleozoic era, primarily from lava materials. The saltiness is a product of acidic gases from the volcanoes (hydrochloric, sulfuric, and carbonic acids) acting to leach ions out of the common silicate rocks. Paleontological studies indicate the change in ions must have been extremely slow, as demonstrated by the narrow tolerance of organisms then alive, such as corals, echinoderms, brachiopods, and radiolarians. Present river ion concentrations differ drastically from the ocean's values, however, indicating a different atmospheric environment in the past. Robert M. Garrels and Fred T. Mackenzie have divided ocean history into three periods. In the earliest, until 3.5 billion years ago, water and volcanic acid emissions actively attacked the crust, leaching out ions and leaving residues of alumina and silicates. The next period, from 3.5 to 1.5 billion years ago, saw slow continuous chemical action attacking the sedimentary rocks, adding silica and ferrous ions. Period three, from 1.5 billion years ago onward, added ions until seawater composition reached apparent equilibrium with a mixture of calcite, potassium-feldspars, illite-montmorillonite clays, and chlorite.

Because the composition of ocean water has remained similar over much of geologic time, generally output must equal input of ions, so geochemical "sinks" must balance geochemical sources. Calcite (calcium carbonate) and silica (silicon dioxide) are removed by marine organisms to form skeletons and shells. Metals are dropped from seawater as newly formed mineral clays, oxides, sulfides, zeolites, and as alteration products at the hot-water basaltic ridges. Sulfur is removed as heavy-metal sulfides precipitating in anaerobic environments, while salts are moved in pore waters trapped in sediments. Residence times for many of the ions have been determined: for example, sodium cycles in 210 million years, magnesium in 22 million, calcium in 1 million, and silicon in 40,000. With such effective removal systems, it is truly a measure of the geochemical resistivity of the Earth's oceans to change that allows the composition to remain so stable for 4 billion years.

METHODS OF STUDY

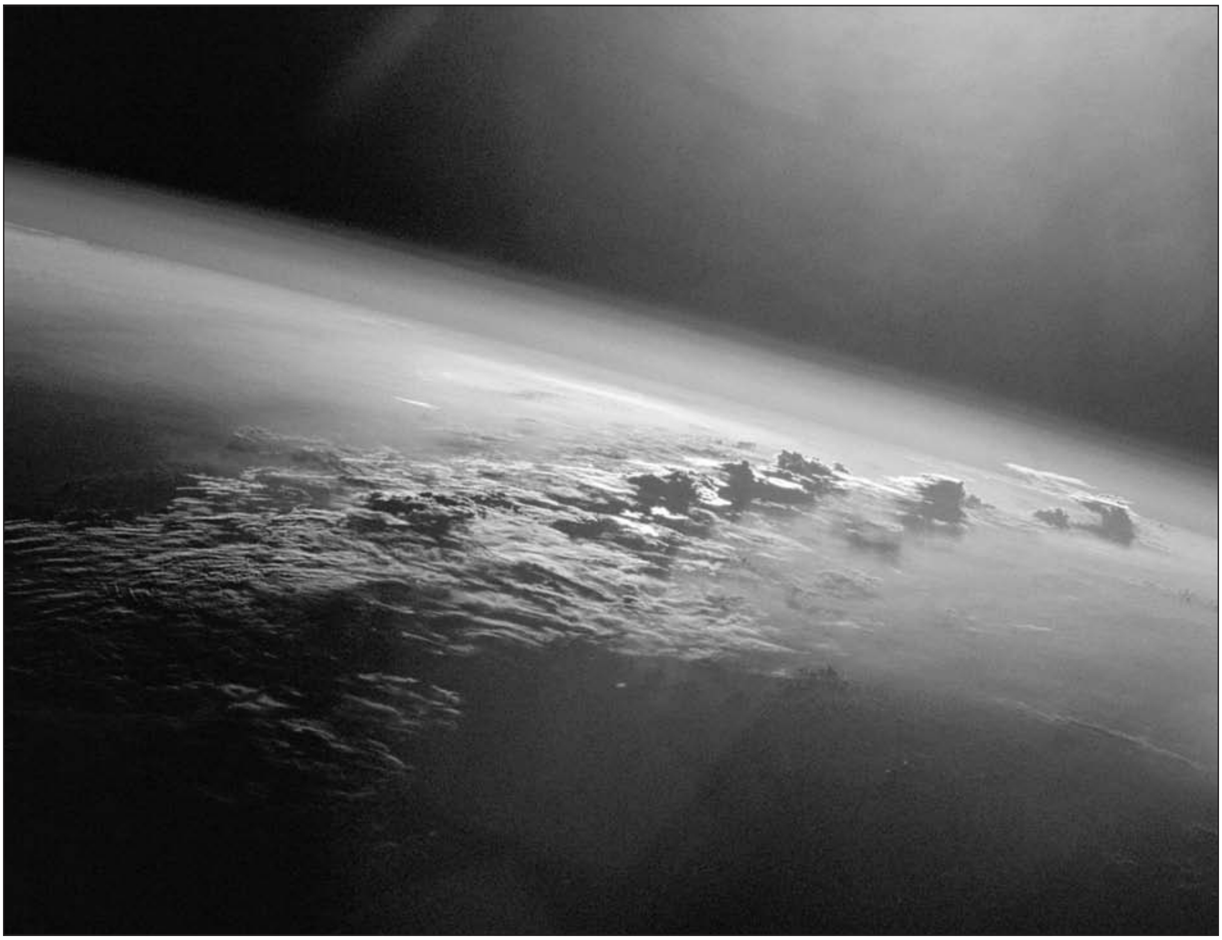
Numerous avenues of approach have been used to investigate the ocean and its ions, including geological, chemical, and physical means. Geology has supplied basic data on the types and makeup of rocks from the earliest solidified materials to present depositional formations. Use of the petrographic microscope, involving thin sections of rocks seen under polarized light, allows the identification of minerals, providing quantity measurements of water attached to the minerals themselves. Paleontological studies of fossil organisms and paleosoils indicate the range of ions in the sea at diverse geologic periods, both by the ions themselves left in the deposited soils and rocks and through studies of the tolerance ranges for similar, twentieth century organisms. Such studies—along with sedimentological investigations of rates and types of river depositions, dissolved ion concentrations, and runoff rates for falling rain—provide determinants for comparing ion concentrations with those in the past for continentally derived materials.

Chemical analysis reveals the various ions present in seawater and rocks via two principal methods. The mass spectrometer identifies types and quantities of ions present by use of a magnetic field to accelerate the charged ions along curved paths, the curvature of the paths based on the weight and charge of the ions. Collection at the end of the paths provides pure samples of the different ions present. For solid samples, electron beam probes analyze an area only one micron in diameter. The electrons, fired at the sample, cause characteristic X rays to be emitted. The energies of the X rays identify the elements or compounds present in the sample.

Solubility studies provide residence times. Similar laboratory projects, testing the ability of water to dissolve and hold ions in solution, argue for a primordial Earth atmosphere that was essentially neutral or mildly reducing in nature. Such reduction characteristics are based on studies of the composition of Earth, supplemented by the composition of Venus and Mars as revealed by various "lander missions."

Missions in interplanetary space have also provided chemical compositions for meteoritic gases, cometary tails and nuclei, and the mixing ratios for noble gases, important for determining the origin of the solar system. Analysis of radioactive isotopes such as helium 3, an isotope of mantle origin, has allowed geophysicists to treat the Earth's mantle as a major elemental source and sink for the various geochemical cycles.

Laboratory analysis reaches two other areas. Petrographic studies of returned lunar rocks reveal that the Moon is devoid of water, lacking even hydroxyl ions. This discovery helps eliminate the solar wind and meteoritic impact as major factors in forming Earth's oceans. Furthermore, high-temperature, high-pressure metallurgical and chemical studies indicate that molten granite, at temperatures of 900° Celsius (1200 kelvins) and under 1,000 atmo-



Earth's Indian Ocean, from the space shuttle Discovery *during the STS-96 mission in 1999.* (NASA) spheres of pressure, will hold 6 percent water by weight, while basalt holds 4 percent. Based on geochemical calculations of the amounts of magma in the planet and lavas extruded over the first billion years, all the ocean's waters can be accounted for, particularly if parts of the fluid, as steam under pressure, are a result of oxidation of deep-seated hydrogen deposits trapped within or combined with mantle rocks.

CONTEXT

Water is a ubiquitous and by far the most important molecule on Earth. All living organisms require it as a basic component of cellular structure and for numerous functions inside the body. The origin of Earth's water is highly significant, because the very presence of water may have set the scheme for all subsequent evolution, both geological and biological, on the planet. During the formation of the solar system, the accretion of various materials trapped water by hydration. Tied to the minerals, the water molecules were released through outgassing by volcanoes and other vents acting as pressure escape valves for the molten interior of the Earth. The water and other volatile gases that were released formed the atmosphere and subsequent oceans. A vital interchange was established between the ground and the atmosphere, one replenishing elements and compounds as they were lost through geochemical sinks in the normal course of history. Water, at first in the atmosphere, then as liquid seas, apparently helped to mediate the greenhouse effect, a mechanism which, if allowed to act unhindered, would have trapped infrared radiation from the Sun and overheated the early Earth. Such actions would have given the Earth the characteristics of the planet Venus: enormously hot and totally inhospitable for life's occurrence.

The outgassed water, settling as rain, also played the dominant role in shaping the landforms of Earth. As a mechanism for fluidization of rocks, it controls to a large extent the motions of magmas, helping them rise to the surface. As a weathering agent, water, in the forms of rain, snow, and ice, carves away the landscape, removing elements, as ions, to the sea. In that location, these elements became usable by early organisms for fulfilling their biological needs, such as home building or metabolism. Water acts as a transport mechanism, a mixing agent, and ultimately a removal tool for maintaining a delicate ionic concentration range within the ocean itself. Evaporating seawater, falling as rain, breaks up rocks and forms soils with nutrients available for land-based plant life, and it provides the freshwater so necessary to non-ocean-dwelling organisms. Without the initial interplay of water on Earth, our planet, instead of being the home of countless billions of creatures, would undoubtedly be a desolate ball, revolving forever around the Sun as an improbable abode of life.

Arthur L. Alt

FURTHER READING

Brancazio, Peter J., ed. *The Origin and Evolution of Atmospheres and Oceans*. New York: John Wiley & Sons, 1964. This work is a collection of papers dealing with the chemical problems relevant to the early formation of the fluid parts of the Earth. Tracing all the basic arguments, the criteria for water formation is clearly explained and its relationship to minerals and rocks elucidated. Some heavy reading, charts, extra references.

Chamberlain, Joseph W. *Theory of Planetary Atmospheres: An Introduction to Their Physics and Chemistry*. New York: Academic Press, 1978. A detailed analysis of the characteristics of diverse atmospheres in the solar system, including water contents. By comparisons of chemical compositions and meteorological observations, criteria are established for examining the possible origins for atmospheric gases and oceans. Some mathematics, heavy reading, numerous charts, comprehensive references.

Consolmagno, Guy. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice-Hall, 1994. An introduction to planetary science for beginners and undergraduates. Covers all planets in the solar system.

Frakes, L. A. *Climates Throughout Geologic Time*. New York: Elsevier, 1980. A well-written explanation of how the interaction of the Earth's atmosphere and oceans has caused the climate of the Earth to change over the history of the planet. Beginning with the possible origin of ocean and atmosphere, changes are traced as revealed through the geological and paleontological records.

Garrison, Tom. *Oceanography: An Invitation to Marine Science*. Florence, Ky.: Brooks/Cole, 2007. Explores all aspects of Earth's oceans. Looks at how humans use and study the oceans. Covers events such as the December, 2004, earthquake and resulting tsunami and Hurricane Katrina.

Henderson-Sellers, A. *The Origin and Evolution of Planetary Atmospheres*. Bristol, England: Adam Hilger, 1983. This work details the theories of where the volatiles for the Earth came from and how the oceans and other planetary atmospheres came into existence from the creation of the solar system. Advanced reading.

Holland, Heinrich D. *The Chemistry of the Atmosphere and Oceans*. New York: John Wiley & Sons, 1978. A detailed reference on the basic chemical elements present in the two media and the wide variety of reactions occurring in each area. The interactions of the two areas are stressed, as are their common origin from materials outgassed from within the Earth. The action of their chemicals on the terrestrial areas is described in detail. Contains references and numerous charts of data. Advanced reading.

Knauss, John. *Introduction to Physical Oceanography*. 2d ed. Long Grove, Ill.: Waveland Press, 2005. Designed for undergraduates, covers a broad range of topics. Also includes material helpful for physicists, geologists, biologists, and chemists. Contains both descriptive and mathematical explanations.

McElhinny, M. W., ed. *The Earth: Its Origin, Structure, and Evolution*. New York: Academic Press, 1979. A readable work dealing with all the basic elements of Earth science. Starting with the theory of planetary formation, the volume covers the origin of the oceans, atmosphere, land, and life-forms. Offers an excellent description of the changes occurring on the planet throughout geologic time.

Pickard, George L. *Descriptive Physical Oceanography: An Introduction*. 5th ed. New York: Pergamon Press, 1990. An introductory work geared toward undergraduates. Covers all aspects of oceanography, including sea-ice physics, thermohaline circulation, and coral reefs.

Ponnamperuma, C., ed. *Cosmochemistry and the Origins of Life*. Dordrecht, the Netherlands: Reidel, 1982. A collection of works dealing with the distribution of elements in the universe, particularly those necessary for life. Provides information on the formation of the planetary system, showing how the chemicals combined at various temperatures to make the planets as different as they are. Discusses origins of oceans, atmospheres, and life; detailed reading with many charts and an extensive bibliography.

Seibold, E., and W. Berger. *The Sea Floor*. New York: Springer, 1982. A delightful book covering the chemistry, geology, and biology of the bottom of the Earth's oceans. That the oceans are a result of outgassing is emphasized. Well written, with a very interesting section on the black and white smokers and their relation to the origin of waters and life.

Trujillo, Alan, and Harold Thurman. *Essentials of Oceanography*. 9th ed. Upper Saddle River, N.J.: Prentice Hall, 2007. An introductory work designed to explain the complexities of oceanography to the average reader. Covers tsunamis and Hurricanes Katrina, Rita, and Wilma.

See also: Auroras; Comets; Earth-Moon Relations; Earth-Sun Relations; Earth System Science; Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Mantle; Earth's Origin; Earth's Rotation; Earth's Shape; Earth's Structure; Eclipses; Greenhouse Effect; Van Allen Radiation Belts.

Earth's Origin

Categories: Earth; The Solar System as a Whole

The Earth's early formation, its subsequent internal differentiation, its active plate tectonics, and its external weathering have left little substantive evidence of its origin intact for direct study. Much about the materials and formative processes involved in the planet's origin can be deduced, however, from seismology, geomagnetics, and the study of meteorites and comets.

OVERVIEW

In order to understand the origins of the Earth, it is necessary to be aware of the sources of the materials from which it is made. The matter from which the Earth and the entire universe is made was created in the big bang, about 13 to 14 billion years ago. The processes that occurred in the first few minutes after the big bang produced all the hydrogen and most of the helium in the universe today; trace amounts of lithium and beryllium also were formed.

Later, after stars formed, they produced all the other chemical elements through nuclear fusion reactions, also called nucleosynthesis. For most of their lives, stars generate the energy to shine by fusing lighter atomic nuclei together to make heavier atomic nuclei. This requires high temperatures and densities, the conditions that exist in the interiors of stars. The first step is the fusion of four hydrogen nuclei into one helium nucleus. The next step is the fusion of three helium nuclei into one carbon nucleus, maybe adding a fourth helium to produce oxygen. This is the end of energy generation and nucleosynthesis for a Sun-like star before it puffs off its outer layers as a planetary nebula, and the exposed core cools and fades to end its life as a white dwarf and ultimately as a black dwarf. (Note that planetary nebulae have nothing to do with planets or the formation of planets. The name dates back to the 1700's, when, viewed through telescopes of that time, they looked fuzzy, like nebulae, and round, like planets).

More massive stars have a more spectacular demise. After carbon and oxygen have formed, further fusion reactions continue to form heavier elements up to iron. The production of elements even heavier than iron does not generate energy but requires the input of energy. The iron core collapses, and the outer layers collapse on top of it and then rebound, tearing the star apart in a supernova explosion. The tremendous energy released in a supernova permits the formation of the rest of the chemical elements. The chemical elements produced during the star's life and death are dispersed by the supernova explosion into interstellar space, there to enrich clouds of gas called nebulae (containing mostly hydrogen and some helium) in the heavier chemical elements.

The Sun, planets, and other bodies of the solar system formed as the result of the gravitational contraction of part of such a nebula about 4.5 to 4.6 billion years ago. The portion that would become the solar system, called the solar nebula, initially was perhaps about a light-year (about 9.5 trillion kilometers) across and was composed of about 74 percent hydrogen, about 24 percent helium, and about 2 percent all the other chemical elements. The exact mechanism responsible for the initiation of the solar nebula's contraction is still speculative. It may have involved the compression of the nebula as it passed through a spiral arm density wave as the nebula orbited the center of our galaxy, the Milky Way. It may have been triggered by a shock wave propagating through the nebula when a nearby massive star went supernova. It is generally agreed that once started, gravitational effects within the solar nebula kept the process going.

Any initial slow rotation of the solar nebula increased its speed with the contraction of the nebula to conserve angular momentum. (The same effect is seen on spinning figure skaters, whose rotational speed increases as they bring their arms close to their bodies.) The increase in rotational speed caused the nebula first to become oblate and eventually to form a flattened equatorial disk. Most of the solar nebula's mass concentrated at the center of the disk, forming the proto-Sun, which grew hotter by gravitational contraction.

As the proto-Sun formed at the center, fractional condensation began—a process in which gaseous matter solidifies into small, sand-sized grains only in regions where the ambient temperature is below the material's melting point. Only metallic grains of iron and nickel condensed close to the proto-Sun. Farther out, where the temperature was lower, they were joined by grains of silicate minerals. Still farther out, various ices of water, carbon dioxide, methane, and ammonia could condense. These solid grains collided with one another and stuck together in a process called accretion, forming planetesimals that grew in size. Within a time span of a few tens of millions to perhaps one hundred million years, the largest planetesimals grew into protoplanets, while the smaller ones became the many satellites and other minor members of the solar system.

As the protoplanets continued to grow, their gravitational influence grew as well. They could attract greater amounts of disk material, thus accelerating their growth while at the same time sweeping the surrounding interplanetary space clean. Solar radiation could then penetrate the space between the Sun and the planets, bringing light and heat to their still-evolving surfaces. It was during this time that the planets started to evolve in different ways. The third planet from the Sun, Earth, and its inner solar system neighbors (Mercury, Venus, and Mars) had relatively weak gravitational fields, which, coupled with their now high surface temperatures and exposure to the solar wind, caused them to lose significant amounts of the lighter gases. This first atmosphere, probably consisting of hydrogen, helium, methane, ammonia, carbon dioxide, and water vapor (gases common in the solar nebula), escaped from the inner planets and was blown away into the outer solar system. The outer planets, Jupiter and beyond—because of their colder temperatures, their greater masses, and consequently the lessened influence of the solar wind—were not so affected. As a result, they became the low-density gas/liquid/ice "giant" planets with small, rocky/metallic cores.

As the accretion process drew to a close, the Earth (along with the other inner planets) was subjected to a final intense bombardment of impacting planetesimals. As each colliding object struck the Earth's surface, its energy of motion was converted into heat energy. Furthermore, radioactivity levels were much higher in the very early Earth, since many of the radioactive elements with shorter half-lives had not yet decayed. As the Earth grew larger in size, it tended to insulate itself, making it more difficult for the energy released by radioactive decay in its interior to reach its surface and escape. All of these effects served to increase the early Earth's temperature to the point that it at least partially melted, allowing chemical differentiation and the development of the Earth's layered internal structure. Molten blobs of heavy metals like iron and nickel sank to the center, forming the iron-rich core. Less dense silicate and oxide minerals remained behind, forming the mantle. The surface also melted, forming a magma ocean perhaps a few hundred kilometers deep. Eventually the surface cooled and hardened into a thin, primitive basaltic crust, probably similar to present-day ocean-floor crust.

At the same time, outgassing released gases trapped in the interior at a prodigious pace, producing the Earth's second atmosphere, probably consisting mostly of water vapor, carbon dioxide, and sulfur dioxide, with smaller amounts of nitrogen, hydrogen sulfide, and other gases, but no free oxygen. (Outgassing continues today through volcanoes, fissures, and fumaroles, but at a much reduced rate.)

As the Earth cooled, water vapor in the atmosphere condensed, formed clouds, and fell as rain, forming the first streams and oceans. Due to the abundant carbon dioxide along with sulfur dioxide and hydrogen sulfide in the atmosphere, this early rain was highly acidic, resulting in rapid chemical weathering of surface rocks. The weathering products were carried by streams into the oceans, rapidly increasing their salinity. By about 4 billion years ago, the oceans had reached nearly their present volume and degree of saltiness. Large amounts of carbon dioxide from the atmosphere dissolved in the oceans, combined with other dissolved materials, and precipitated out as sediment (mostly as the mineral calcite, as calcium carbonate). As other gases were removed from the atmosphere, nitrogen remained, eventually becoming the major atmospheric constituent.

Probably by about 3.8 billion years ago, the first life appeared. Organic molecules, including amino acids, may have formed in the Earth's early atmosphere and oceans with solar ultraviolet light, lightning, or deep-sea hydrothermal vents providing the needed energy input, or they may have been delivered to the Earth's surface by impacts of comets, asteroids, and meteoroids. The earliest living organisms were anaerobic (able to survive without oxygen), since there was no free oxygen in the atmosphere and oceans. Then cyanobacteria and possibly other early organisms developed photosynthesis, using sunlight to turn water and carbon dioxide into sugars for food. This reaction released free oxygen into the oceans and atmosphere, and life evolved to utilize it to extract energy from food.

Thus the Earth was transformed from its formative stages to what it is today, a place where life thrives, where rocks are formed and weathered on the surface, and where a dynamic interior drives tectonic processes.

METHODS OF STUDY

Much of what is known about the origins of the Earth is derived by studying meteorites and comets as well as from seismology and geomagnetics. Meteorites and comets are unaltered or little-altered examples of early solar-system materials, providing information on the composition of and processes that occurred in the early solar nebula and the various types of bodies that formed from it. Seismology and geomagnetics give researchers clues about the internal structure of the planet.

Meteorites are extraterrestrial pieces of rock or metal that survived their fall through the Earth's atmosphere. The combined total composition of all meteorites is probably representative of the rocky and metallic material from which the inner planets—Mercury, Venus, Earth, and Mars—formed. Comets provide evidence of the more volatile components of the early solar system. They are composed of various "dirty" ices, indicative of materials blown out from the inner solar system by the solar wind but not before a portion was incorporated into the accreting planetesimals.

Data obtained from seismology (the study of the transmission of earthquake shock waves through the Earth) led to the discovery that the Earth's interior is divided into several distinct layers or zones. Observations of a change in speed of seismic waves near the Earth's surface led to the discovery of the Mohorovičić (Moho) discontinuity, the boundary between the crust and mantle. A low seismic velocity zone is now recognized in the upper mantle below the Moho and is used to define the lower boundary of the Earth's rigid lithosphere and the top of the "plastic," deformable asthenosphere. Another seismic discontinuity 2,900 kilometers below the surface delineates where the solid mantle is separated from the molten outer core. Later seismic work revealed the existence of a solid inner core.

Studies of the magnetic field of the Earth also give support to the zonal nature of the planet's interior. Hypotheses concerning the generation of the magnetic field within the Earth assume an iron-nickel-rich core (not unlike the iron-nickel meteorites) with a solid interior surrounded by a molten outer part. This combination could produce an electric dynamo that could sustain a magnetic field.

All these disciplines provide evidence of the Earth's formation in the early solar nebula. As more data are obtained, the picture of the Earth's origin becomes clearer and more refined.

CONTEXT

An understanding of the Earth's origin has many practical benefits. For example, the genesis of ore bodies is invaluable to prospecting for new resources. By knowing the products of various processes in the past, one is better able to predict human impact on present environments. Planetary engineering can use such information for the modification or preservation of conditions on the Earth, and maybe someday on other planets such as Mars. Meteorite size and shape studies were employed by space engineers in designing reentry vehicles and in studying their aerodynamic properties. Theories about material behavior in zero-gravity conditions similar to those in the solar nebula have led to experiments on manufacturing techniques in Earth orbit that are impossible to conduct on the Earth's surface.

Bruce D. Dod

FURTHER READING

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. A general overview of the solar system and its components, this book has been organized around comparative planetology. A discussion of the various aspects of the genesis of the planets is scattered throughout. Draws heavily on the results of space exploration and on the interdisciplinary use of science to illustrate many concepts. Contains abundant illustrations, such as full-color photographs, artwork, graphs, and charts.

Brush, Stephen G. *Nebulous Earth: The Origin of the Solar System and the Core of the Earth from Laplace to Jeffreys*. New York: Cambridge University Press, 1996. An in-depth reference work detailing twentieth century theories on solar-system formation. Also discusses lunar origin theories.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A well-written college-level textbook for introductory astronomy courses. Offers an entire chapter on the formation of planets.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Several sections deal with the origin of the solar system.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook, thorough and well written. Part of one chapter deals with the origin of the solar system, and part of another specifically addresses the development of the Earth.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college textbook beyond the introductory level, with an approach based on comparative planetology. Offers much information on the origin of the solar system in general and the Earth in particular.

Horton, E., and John H. Jones, eds. *Origin of the Earth*. New York: Oxford University Press, 1990. A collection of articles written by experts in the field. Focuses on the study of the Earth and origins of the Earth-Moon system.

Hutchison, Robert. *The Search for Our Beginning*. New York: Oxford University Press, 1983. The author addresses the problem of determining the processes involved in the formation of the Earth and other solar-system bodies through the analysis of meteorites. Links astrophysics, geology, cosmochemistry, organic chemistry, and astronomy using meteoritics as the common ground. Provides historical perspectives along with space exploration results. Contains some fine illustrations, both in color and in black and white.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. Designed as a text for a college course in planetology, this book contains many references to the origins of the solar system and its individual components. Comparative planetology based on space exploration results, meteoritics, and other sources are utilized throughout the text to illustrate some of the evolutionary phases in the development of the planets and other solar-system objects. Extensively illustrated.

Ozima, Minoru. *The Earth: Its Birth and Growth*. Translated by J. F. Wakabayashi. New York: Cambridge University Press, 1981. Traces the genesis of the Earth and its growth while highlighting problems addressed by isotope geochemistry. The past 4.5 billion years are sketched in terms that are easy to comprehend. Alternative hypotheses and explanations are considered.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A thorough college textbook for introductory astronomy courses, divided into many short sections on specific topics. An entire unit is devoted to the origin of the solar system.

Smart, William M. *The Origin of the Earth*. 2d ed. New York: Cambridge University Press, 1953. An older work useful for those interested in earlier hypotheses and theories of the Earth's origin.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and fully illustrated. Contains a chapter on the origin and historical development of the Earth.

Wasson, John T. *Meteorites: Their Record of Early Solar-System History*. New York: W. H. Freeman, 1985. Written as a text for a course on solar-system genesis, this book includes topics on meteorite classification, properties, formation, and compositional evidence linking meteorite groups with individual planets. Describes how researchers use meteorites to determine conditions in the formative periods of Earth and other planets. Includes many graphs, charts, and illustrations.

See also: Comets; Earth-Moon Relations; Earth-Sun Relations; Earth System Science; Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Mantle; Earth's Oceans; Earth's Rotation; Earth's Shape; Earth's Structure; Lunar History; Solar System: Element Distribution.

Earth's Rotation

Category: Earth

The rotation of the Earth results in the days and nights that provide the daily rhythm of life. Rotation causes the Earth to be flattened at the poles and to bulge at the equator. It also produces the Coriolis force, which influences the circulation of the atmosphere and oceans.

OVERVIEW

The spinning of the Earth on its polar axis is called rotation. The ancient Greeks considered the Earth to be a motionless body in the center of a geocentric (Earth-centered) universe. An exception was Heracleides (fourth century B.C.E.), who thought that the Earth did rotate. In general, the Greeks reasoned that, in their experience, if the Earth moved, they would feel some effects of it. Later, the work of Nicolaus Copernicus, Galileo Galilei, Johannes Kepler, and Sir Isaac Newton resulted in the paradigm shift to a heliocentric (Sun-centered) system in which the Earth was a planet simultaneously rotating on an axis while revolving around the Sun. The Earth rotates from west to east, thus making the Sun, Moon, planets, and stars appear to move from east to west across the sky. We commonly refer to the Sun, Moon, planets, and stars as rising and setting because it appears that they all are moving around the Earth, when in fact the Earth is rotating on its axis while revolving around the Sun.

The Earth's axis of rotation is inclined 23.5° from the perpendicular to the ecliptic (the plane of the Earth's orbit around the Sun); thus the axis of the Earth makes an angle of 66.5° to the plane of the ecliptic. The inclination causes the seasons (and the seasonal variation in the length of day and night) as the Earth orbits the Sun during the course of a year. When either the Northern or Southern Hemisphere of Earth is tilted toward the Sun, the period of daylight is longer and night is shorter in that hemisphere. The longer duration of daylight, coupled with the Sun's rays striking that hemisphere more nearly head-on, results in summer in that hemisphere. Conversely, when either hemisphere is tilted away from the Sun, the duration of daylight is shorter and night is longer, the Sun's rays strike that hemisphere more obliquely, and that hemisphere experiences winter.

Because of rotation, a point on the Earth's equator moves 1,674 kilometers per hour; the speed decreases to 1,450 kilometers per hour at 30° north or south latitude, and to 837 kilometers per hour at 60° north or south latitude. The Earth's rotation defines the unit of time called the "day." A day is defined as the interval of time between successive passages of a meridian, or line of longitude from the North to South Pole, under a reference object (for example, the Sun or a star). A day with reference to the Sun is called a solar day, and a day with reference to the stars is called a sidereal day. Because of the Earth's orbital motion around the Sun in one year, the Sun appears to move eastward relative to the stars approximately one degree per day. This makes the solar day approximately four minutes longer than the sidereal day, since the Earth must rotate a little bit farther to complete one rotation relative to the Sun as compared to the stars.

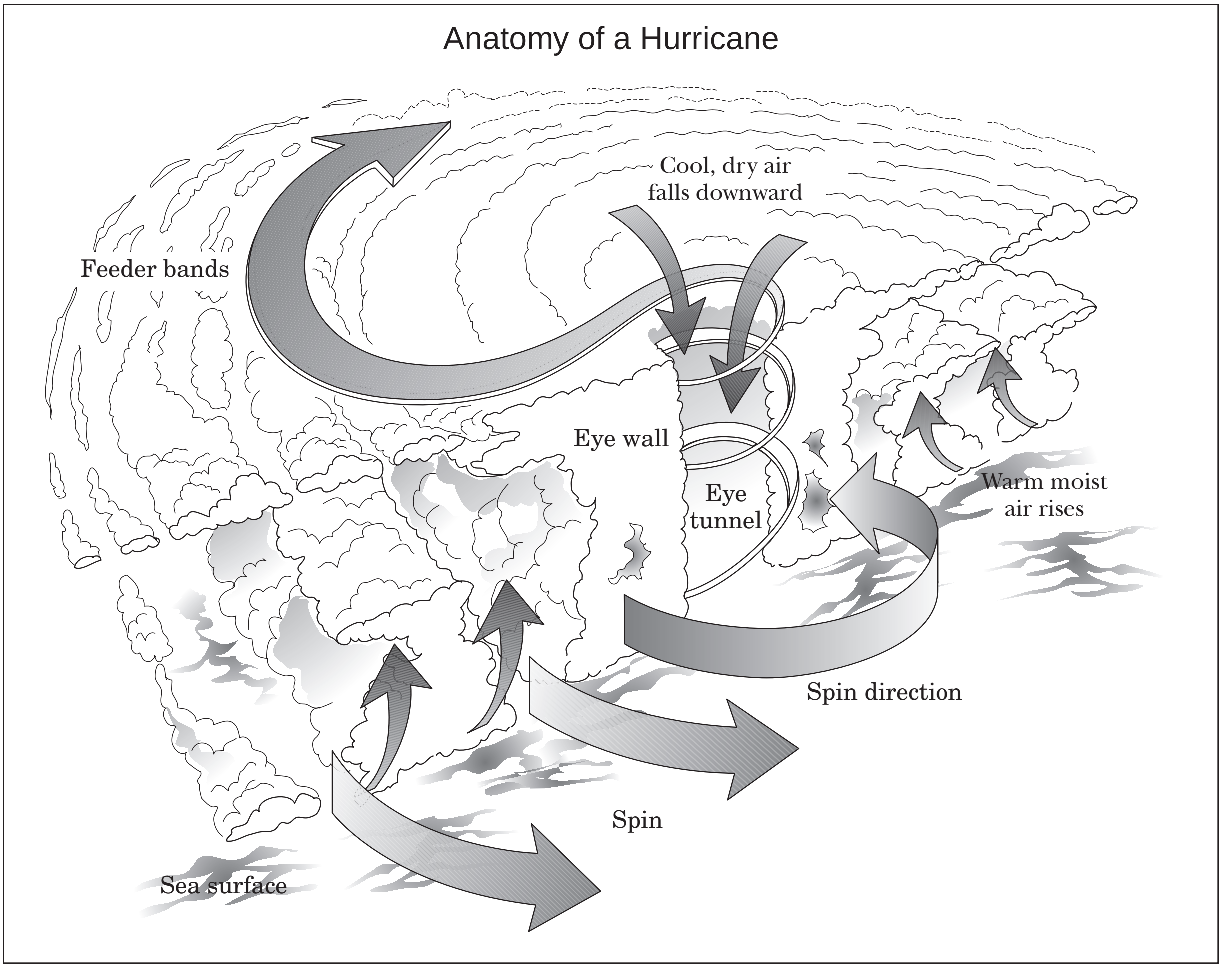
The Earth's orbit around the Sun is slightly elliptical, and the Earth's orbital speed varies with its distance from the Sun, being fastest when Earth is closest to the Sun (perihelion), around January 3, and slowest when Earth is farthest from the Sun (aphelion), around July 4. This means the Sun's apparent motion relative to the stars varies during the year, being greatest when the Earth's orbital speed is fastest at perihelion and smallest when the Earth's orbital speed is slowest at aphelion. Thus the length of the solar day as measured by a sundial (called the apparent solar day) varies slightly during the year. The length of the apparent solar day averaged over a year is called the mean solar day, and mean solar time is the basis for the 24-hour day (of 86,400 seconds) kept by clocks. (The sidereal day is 23 hours, 56 minutes, and 4.091 seconds long.)

Because of rotation, the Earth is flattened in the polar regions and bulges at the equator, thus making it slightly ellipsoidal, an oblate spheroid. The equatorial radius is 6,378 kilometers, while the polar radius is 6,357 kilometers. Thus a point on the equator is 21 kilometers farther from the center of the Earth than either pole is. Because the Earth is slightly flattened, the length of a degree of latitude changes from 109.92 kilometers at the equator to 111.04 kilometers at the poles. Also, because a point on the equator is farther from the center of the Earth and its rotational speed is faster, the effect of gravity is reduced there compared to other points on Earth. Thus an object at the equator weighs less, about 1 pound in 200, compared to the same object at either pole.

The Coriolis force or effect, named after a nineteenth century French engineer who studied this phenomenon, is caused by the Earth's rotation. It is an apparent force that affects free-moving bodies (such as wind, water, or missiles). For example, a fired projectile will veer to the right relative to the Earth's surface in the Northern Hemisphere, and to the left in the Southern Hemisphere. Precisely on the equator itself, free-moving objects are not deflected, but as they move north or south of the equator, the deflection becomes more pronounced.

The Coriolis effect is responsible for the global prevailing wind belts. Warm air near the equator rises and flows toward the poles. Cooled at higher altitude, the air descends around 30° north and south latitudes and spreads out both toward the equator and the poles. Air flowing toward the equator is deflected westward in both hemispheres, producing the easterly trade winds of the tropics. Air flowing toward the poles is deflected eastward in both hemispheres, producing the westerly winds of temperate latitudes. The Coriolis effect determines the direction wind blows around local high and low pressure systems in the atmosphere. Air moves outward from high pressure systems and inward toward low pressure systems. In the northern hemisphere, the moving air veers toward the right, setting up clockwise rotation around atmospheric highs and counterclockwise rotation around atmospheric lows. In the southern hemisphere, the moving air veers left, setting up counterclockwise rotation around highs and clockwise rotation around lows. This effect is especially noticeable in hurricanes (also called typhoons), which are regions of extremely low atmospheric pressure. The Coriolis force also influences ocean currents, which in turn affect the climate of coasts they flow along. The ocean currents of the Northern Hemisphere tend to flow clockwise, while those of the Southern Hemisphere flow counterclockwise. Witness the Gulf Stream of the north Atlantic and the Japanese (Alaskan) current of the north Pacific, always turning to the right, while the south Atlantic flow is to the left.

Overall, the rotation of the Earth is slowing down, and the length of the day is increasing by milliseconds per century. The decrease in rotational speed is primarily a result of the tidal fric-



The Coriolis force, which influences the circulation of the atmosphere and oceans, is a result of Earth's rotation and determines the direction wind blows around local high and low pressure systems in the atmosphere. The effect is especially noticeable in hurricanes, regions of extremely low atmospheric pressure.

tion caused by the gravitational pull of the Moon and to a lesser extent the Sun. Evidence for a lengthening day during geologic history comes from the study of fossils. Clams, corals, and some other marine invertebrates add a microscopically thin layer of new shell material each day, and the thickness varies seasonally throughout the year. Counting the daily growth lines in an annual set in well-preserved fossils yields the number of days in a year. Since the length of a year (the period of the Earth's orbit around the Sun) presumably has not changed, the length of a day in past geologic times can be determined. During the early Cambrian period (540 million years ago), there were 424 days in a year, and thus each day was about 20 hours, 40 minutes long. In the late Devonian period (365 million years ago), a year consisted of 410 days, each about 21 hours, 23 minutes long. At the beginning of the Permian period (290 million years ago), a year was down to 390 days, each about 22 hours, 29 minutes long.

As Earth's rotation slows in response to the Moon's tidal drag, the Earth's rotational angular momentum is transferred to the Moon, which increases the Moon's orbital angular momentum around the Earth, causing the Moon to move outward, away from Earth. This in turn increases the Moon's orbital period around Earth. This will continue until Earth's rotation on its axis is tidally synchronized with the Moon's revolution around Earth, both Earth and Moon keeping the same side facing each other. To conserve angular momentum, the Moon's distance and orbital period will increase to 549,000 kilometers (341,000 miles) and 46.7 of Earth's present days. Thus Earth's sidereal rotation period will be 46.7 days, and the mean solar day (time from noon to noon) will be 53.5 of Earth's present days. Since the length of the year will be unaffected, there will be only 6.8 solar days in a year.

Currently Earth's day is lengthening by an average of about 2×10^{-5} seconds each year, but the rate is quite erratic and sometimes even speeds up a bit. If this current average slowdown rate is extrapolated into the future, it will take 2×10^{11} years for Earth and Moon to become tidally locked. Alternatively, laser ranging (using retroreflectors left on the Moon's surface by the Apollo Moon landings) shows that the Moon currently is moving away from Earth at 3.8 centimeters per year. Extrapolating this rate into the future, it will take 4×10^{9} years for the Moon to reach its final distance. The disagreement of these two time estimates indicates that the rates of slowdown of Earth's rotation and the increase in the Moon's distance probably will not remain constant, and the time to achieve Earth-Moon tidal lock is probably at least billions of years. However, before this can occur, the Sun probably will expand and become a red giant first.

Numerous detailed studies show that, superimposed on the systematic long-term slowdown, Earth's rotation has numerous small random changes. The reasons for these variations include transfer of angular momentum between different parts of the Earth's interior; transfer of angular momentum between the atmosphere, the oceans, and the Earth's surface; movement of air masses and changes in wind patterns; growth or shrinkage of polar ice caps; volcanic activity; earthquakes; and plate tectonic movements. The timescales for the various effects on the Earth's rotation vary, from short-term to long-term, and from systematic seasonal to erratic. It is evident from many studies that the Earth's rotational speed has varied throughout geologic time and continues to change on a daily, weekly, monthly, seasonal, yearly, and even longer-term basis.

Currently the Earth's axis of rotation points toward Polaris, Earth's present North Star. However, the Earth's axis slowly changes direction in space in a process called precession. Recall the Earth has an equatorial bulge, and it is tilted about 23.5° from the ecliptic, the plane of the Earth's orbit around the Sun. The torque exerted by the gravitational pull of the Moon and the Sun on the Earth's equatorial bulge trying to make it line up with the ecliptic plane causes the Earth's axis to slowly precess, like the axis of a tilted spinning toy top. This precession causes the Earth's axis to trace out in space a double cone (two cones joined at their vertices) with a vertex angle of 47° (twice the 23.5° axial tilt). The Earth's axis slowly shifts direction about 50 arc seconds per year, and it takes about 26,000 years for a complete precession cycle. As the axis points to different parts of the sky in response to precession, stars other than Polaris have served and will serve as north stars. Also, in about 13,000 years (half the precessional cycle), the constellations seen during specific months on Earth will be "shifted" by six months, so that those constellations now seen during June (for example) will be seen during December, and so on. Consequently, the astronomical coordinate system of right ascension and declination slowly and systematically changes during the precessional cycle. Catalogs listing those coordinates for stars, nebulae, galaxies, and other celestial bodies must specify the epoch (year) for which the listed coordinates are rigorously correct. To point a telescope at some desired object some other year requires calculating precessional corrections to the listed coordinates.

Superimposed upon the precessional motion are two other motions. One of these is a small oscillating motion called nutation, which has a semiamplitude of 9.2 seconds of arc and a period of 18.6 years. This motion is associated with the periodic variation in the orientation of the Moon's orbital plane around the Earth with the Earth's orbital plane around the Sun. The other motion, called Chandler's wobble, has two oscillations. One of the oscillations, the Chandler component, with a period of twelve months, is a result of meteorological effects associated with seasonal changes in air masses. The second oscillation of the Chandler wobble, the 14.2-month component, is caused by shifts in the Earth's interior mass. Thus the changing direction of the Earth's rotational axis is not smooth but "wiggly" or "wobbly."

METHODS OF STUDY

Sundials were first used to mark the passage of the apparent solar day, as the shadow of the gnomon (stick) moved across the face of the dial. In 1671, the French astronomer Jean Richer made time measurements with a pendulum clock both in Paris (49° north) and in Cayenne, French Guiana (5° north) and compared the two. In French Guiana, the clock "lost" 2.5 minutes per day compared to Paris. He attributed this loss to a decrease in effective gravitational pull toward the equator due to the Earth's rotation; the practical consequence was that pendulum clocks needed to have the length of their pendula adjusted according to latitude to be able to keep accurate time.

In 1851, the French physicist Jean-Bernard-Léon Foucault hung a 25-kilogram iron ball with a 60-meter-long wire from the dome of the Panthéon in Paris, with a pin at the bottom of the ball to make marks in a smooth layer of sand underneath. After only a few minutes, the tracings in the sand showed that the plane of the ball's swing slowly rotated clockwise as seen from above. Foucault explained this as a demonstration of the Earth's rotation, which moved the attachment point on the dome and the sand on the floor, while the pendulum tried to maintain the plane of its swing in the same direction. In the 1950's, atomic clocks began to be used to measure time accurately over long periods. When time kept by these clocks was compared to time determined by the rotation of the Earth, small variations in the Earth's rotation were found.

Newer techniques used to determine length of day and polar motion involve the use of satellites and lasers. One method, called lunar laser ranging (LLR), involves the emission of light pulses from a laser on Earth to reflectors left on the Moon by Apollo and Soviet spacecraft. The returning pulses of light are received by a telescope. The total travel time is calculated to determine the Earth-to-Moon distance. By observing the time the Moon takes to cross a meridian during successive passages, this method has provided very good length-of-day measurements. Another technique involves the use of the Laser Geodynamics Satellite (Lageos). This satellite is covered by prisms that reflect light from pulsed lasers on Earth. Again, the returned beam is received by a telescope and the round-trip travel time is used to infer the one-way distance from the Earth to the satellite. This method, which includes a network of stations on Earth, can provide insight into yearly movement of crustal plates, which is believed to cause variations in the Earth's rotation. A very accurate technique known as very-long baseline interferometry (VLBI) is also being used to plot continental drift as well as variations in Earth's rotation and the position of the poles. In this method, radio signals from space (typically from quasars) are received by two radio antennas and are tape-recorded. The tapes are compared, and the difference between the arrival times of the signals at the two radio antennas is used to calculate the distance between the two. If the distance between the two antennas has changed, the crustal plates have moved.

CONTEXT

The spinning of the Earth on its polar axis once every twenty-four hours is very much a part of the daily rhythm of life. Among the primary ways Earth's rotation is felt by and governs life are its impact on our day, on gravitational pull, and on the atmosphere. Earth's rotation gives us a daily time reference by the passage of days and nights. The spinning of the Earth on its polar axis causes the Earth to bulge at the equator and to be flattened in the polar areas. Because of this phenomenon, the distance to the center of the Earth varies with latitude, and as a result the effective gravitational pull on objects on the Earth's surface also varies—objects weigh slightly less at the equator than in the polar regions of the world. The Coriolis effect, an apparent force caused by the rotation of the Earth, causes free-moving bodies to be deflected to the right in the Northern Hemisphere and to the left in the Southern Hemisphere. It governs the direction of winds as they flow in or out of pressure systems, establishing the wind belts of the world. The Coriolis force also affects the flow of ocean currents, and these patterns help to alter climates along the coasts of continents.

Roberto Garza

FURTHER READING

Bostrom, Robert C. *Tectonic Consequences of the Earth's Rotation*. New York: Oxford University Press, 2000. Reviews scientific data looking for a link between geotectonics and the Earth's rotation. The author presents a better theory to explain tidal Earth. For the more advanced reader or undergraduate.

Gould, S. G. "Time's Vastness." *Natural History* 88 (April, 1979): 18. This article summarizes the reasons for the slowing down of the Earth's rotation. It discusses the use of corals as a proof that the length of the day is increasing and that the number of days in a year is decreasing. Suitable for high-school-level readers.

Lambeck, Kurt. *The Earth's Variable Rotation: Geophysical Causes and Consequences*. New York: Cambridge University Press, 2005. Focuses on the irregular rotation of the Earth and gives detailed analysis of the various reasons for it. Covers the interdisciplinary fields of solid Earth physics, oceanography, meteorology, and magnetohydrodynamics. Technical.

McDonald, G. E. "The Coriolis Effect." *Scientific American* 186 (1952): 72. The article takes a nontechnical approach to the study of how objects move on the Earth as a result of the Coriolis effect. Suitable for high school readers.

Markowitz, W. "Polar Motion: History and Recent Results." *Sky and Telescope* 52 (August, 1976): 99. This article reviews studies of polar motion, looking at how the Earth's rotation and precessional motions are affected by various forces.

Mulholland, J. D. "The Chandler Wobble." *Natural History* 89 (April, 1980): 134. Discusses how small movements affecting the Earth's axis may be associated with other terrestrial phenomena. Suitable for high school readers.

Munk, W. H., and G. J. F. MacDonald. *The Rotation of Earth: A Geophysical Discussion*. New York: Cambridge University Press, 1960. Dated, but valuable for its detailed analytical treatment of the physics of Earth's rotation. Includes discussion of the small fluctuations in rotation as a result of redistribution of angular momentum, thought to be caused by dynamics in the fluid outer core. Designed for professional geophysicists.

Rosenburg, G. D., and S. K. Runcorn, eds. *Growth Rhythms and the History of the Earth's Rotation*. New York: John Wiley & Sons, 1975. A compilation of studies that can serve as an introduction to the methods of determining the history of the Earth's rotation. The text is suitable for college-level readers not intimidated by technical language. Each study includes a bibliography, and the book is carefully indexed by author, taxonomy, and subject.

Smylie, D. E., and L. Mansinha. "The Rotation of the Earth." *Scientific American* 225 (December, 1971): 80. This article analyzes measurements indicating that the Earth's wobble may be due to earthquakes. It is a well-illustrated article that can be understood by high school readers.

Stephenson, F. Richard. *Historical Eclipses and Earth's Rotation*. New York: Cambridge University Press, 2008. Investigates the history of the Earth's rotation by studying eclipses throughout ancient and medieval times. Shows how tides cannot be solely responsible for the lengthening of the day.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses gives a very clear explanation of the tidal slowdown of the Earth's rotation and specific details on the number of days in the year in past geologic times based on fossil evidence.

See also: Earth-Moon Relations; Earth-Sun Relations; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Shape; Planetary Rotation.

Earth's Shape

Category: Earth

It has been known for centuries that Earth is not a perfect sphere. The diameter of the planet is greater at the equator than it is from pole to pole. This oblateness is the result of Earth's daily rotation on its axis. Even smaller irregularities in shape have been measured by Earth-orbiting satellites.

OVERVIEW

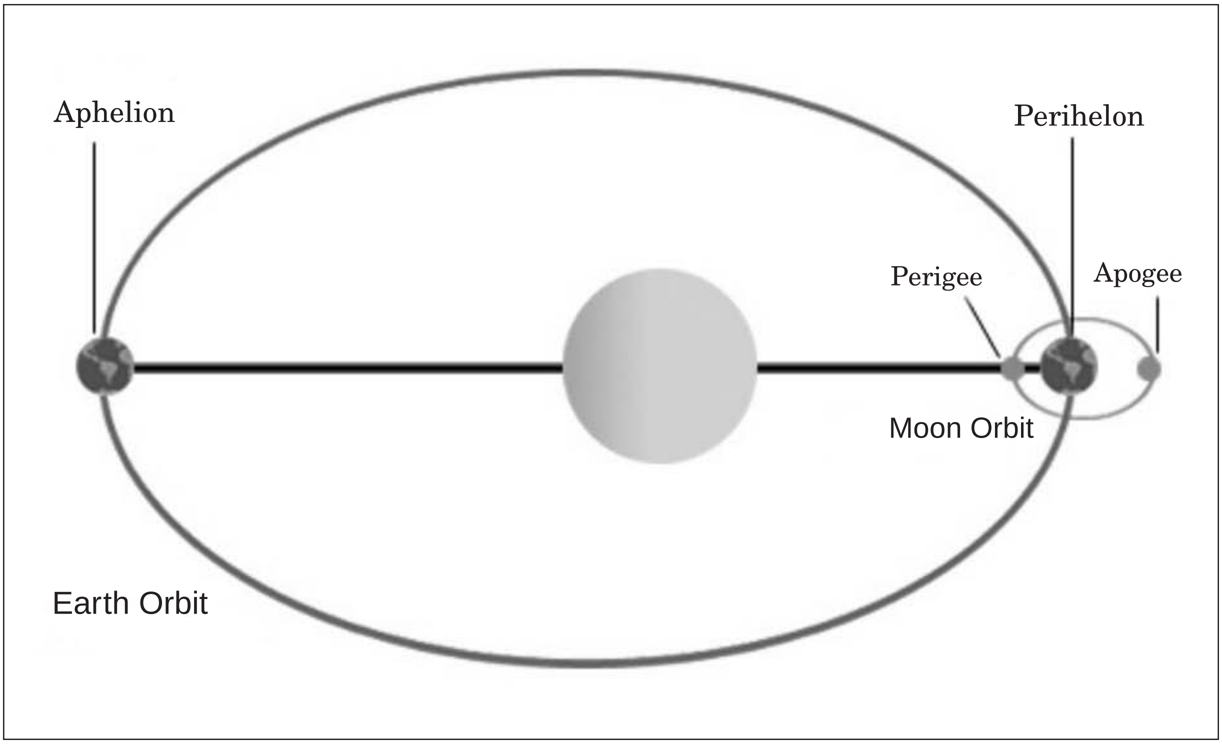
The discovery that Earth is not a perfect sphere dates to the seventeenth century, when measurements of the distance corresponding to one degree of latitude were found to increase systematically from the equator toward both poles. Because of Earth's oblateness, one degree of latitude has a length of 110.6 kilometers at the equator and 111.7 kilometers at the poles.

Besides rotation, other forces imposed upon Earth affect the shape of the planet, a prime example being Earth tides. The oceans of the world generally have two tidal bulges, or regions where the ocean surface is relatively high, caused primarily by the gravitational attraction of the Moon and to a lesser extent by the Sun. When the Sun, Moon, and Earth all are on a straight line (the Sun and Moon either on the same side or opposite sides of the Earth), the oceans display the highest high tides and lowest low tides (spring tides), as the tidal effects of the Sun and Moon reinforce each other. When the line from the Earth to the Moon makes a right angle to the line from the Earth to the Sun, the tidal effects of the Sun and Moon partially cancel out, and the high tides are not very high and the low tides are not very low (neap tides). Not only does the water of the oceans rise and fall because of tides, but so too the surface of the "solid" Earth rises and falls very slightly due to the same tidal forces imposed by the Moon and the Sun. This periodically varying distortion is so slight as to render accurate measurements of it quite difficult.

A view of Earth from satellite distance in space would, to the naked eye, suggest that the planet is a perfect sphere, yet measurements reveal that it is not. All planets (including Earth), along with the larger satellites and asteroids, are essentially spheroidal, while smaller objects are not. Sufficiently small solid objects (such as books, boulders, and bones) can maintain any arbitrary shape because of the strength of the material of which they are composed. However, sufficiently large objects (even if solid) have internal forces due to their self-gravity that are strong enough to overwhelm the strength of whatever material composes them, and they pull themselves into spherical shapes. The critical "threshold" size depends on the density and strength of the solid material, but for the three types of solids (ices, silicate minerals, and metals, mainly iron) most common in the solar system, the threshold is approximately the same: on the order of a few hundred kilometers.

Rotation (spinning on an axis) makes a large object depart from being spherical. The Earth's rotational period, measured in a quasi-inertial frame based on distant stars and galaxies, is 23 hours, 56 minutes, and 4 seconds. As a result of this rotation, a point on the Earth's equator moves with a speed of 1,674 kilometers per hour, a point at 30° north or south latitude moves 1,450 kilometers per hour, and a point at 60° north or south latitude moves 837 kilometers per hour. The increase in rotational speed toward the equator makes the equator bulge outward, transforming a spherical shape into an oblate spheroid. The Earth's equatorial diameter is 12,756 kilometers, while its polar diameter is 12,714 kilometers. Its oblateness is defined as the difference in diameters divided by the equatorial diameter, which gives a value of 0.00336 or about one part in 298.

A comparison of the planets in the solar system shows that rotation plays a dominant role in determining oblateness. Mercury and Venus, each with very slow rotation, have no discernible oblateness and are essentially spherical. Jupiter and Saturn rotate the fastest, and both are noticeably oblate as seen in telescopes and spacecraft images. However, there are other factors. For example, Saturn rotates slightly more slowly than Jupiter, and Mars rotates slightly more slowly than Earth; yet Saturn is more oblate than Jupiter (about one part in 10 compared to one part in 15), and Mars is more



The Earth's and Moon's elliptical orbits—Earth's around the Sun, and the Moon around Earth—affect their tides and hence their shapes. (NOAA)

oblate than Earth (about one part in 200 compared to one part in 300). These "discrepancies" probably are due to the distribution of mass throughout the planet and the rigidity of the material composing the different parts of the planet.

METHODS OF STUDY

Measurements of small departures from an oblate spheroid shape became possible with the advent of the space age and the development of twentieth century instrumentation. Perturbations in the orbits of Earth-orbiting satellites show that the Earth's mass is not distributed as it would be were it a simple oblate spheroid. Some of the earliest data indicated that the Earth is slightly pear-shaped, with small bulges of up to about 100 meters near the North Pole and in a band south of the equator. More recently, satellite-mounted radar altimeters have been able to map continuously the topography of the ocean surface to an accuracy of a few centimeters by bouncing radar signals off the water. With wave crests and troughs averaged out, ocean surfaces show small but significant deviations from a smooth oblate spheroid that reflect the topography of the seafloor underneath. Major seamounts and suboceanic ridges are clearly marked by regions of higher ocean surface above. Likewise, the major deep-sea trenches, as are common around the rim of the Pacific Ocean, are marked by troughs in the ocean surface above. This phenomenon is due to small variations in the acceleration of gravity. In the case of a seamount or ridge, there is a concentration of mass (the rock composing the feature), so gravity there is a bit stronger and attracts more water over it. In contrast, there is a deficit of mass in a trench, so gravity there is slightly weaker and attracts less water over it.

CONTEXT

The passing of geologic time has brought changes in the phenomena that control Earth's shape. The distance from the Earth to the Moon was less than at present, and Earth rotated faster in the past. These two differences would have produced larger and stronger tides, and a larger equatorial bulge. It is interesting to speculate as to what effects those changes might have had on ancient dynamic processes. Today, Earth's inhabitants suffer little effect from the planet's distortion. It cannot be observed with the naked eye, and it does not appear to play a role in weather patterns and climate.

However, one practical, though small, consequence of the Earth's shape is the variation of the effective gravitational acceleration with latitude. Due to Earth's equatorial bulge, which is the result of Earth's daily rotation on its axis, the effective gravitational acceleration, and hence the weight of objects, is slightly less at the equator than at either pole. Because of the equatorial bulge, objects at sea level at the equator are about 21 kilometers farther from the center of Earth than if they were at sea level at either pole. This alone reduces the gravitational acceleration at the equator by about 0.15 percent compared to the value at either pole. The effective gravitational acceleration at the equator is further reduced directly by Earth's rotation, since some of the gravitational acceleration that would otherwise exist is used to provide the centripetal acceleration needed to make objects follow the curved paths that keep them in contact with Earth's surface as Earth rotates. This reduces the effective gravity at the equator by about 0.35 percent. Combining the two effects, the gravitational acceleration varies from 9.832 meters per second squared at either pole to 9.780 meters per second squared at the equator, or about 0.5 percent. As a result, weight at the equator is reduced by about 0.5 percent compared to weight at either pole, so an object weighing 200 pounds at either pole weighs about 1 pound less at the equator.

John W. Foster

FURTHER READING

Greenberg, John L. *The Problem of the Earth's Shape from Newton to Clairaut*. New York: Cambridge University Press, 1995. Covers the early studies to determine the shape of the Earth by various scientists, including Isaac Newton. Explains their influence on Alexis Claude Clairaut, who confirmed Newton's belief that the Earth is flattened at the poles.

Ince, Martin. *The Rough Guide to the Earth 1*. New York: Rough Guides, 2007. A handy reference on most aspects of Earth science. Includes several diagrams and pictures to help explain material. For beginners: nontechnical and easy to read.

James, David E., ed. *The Encyclopedia of Solid Earth Geophysics*. New York: Van Nostrand Reinhold, 1989. A complete reference work on solid-Earth geophysics. Includes more than 150 articles by top scientists. Also covers topics such as geology, seismology, and gravimetry, among others. For advanced readers: detailed and technical.

King-Hele, D. "The Shape of the Earth." *Scientific American* 217 (October, 1967): 17. This article begins with the historical views of the shape of the Earth. It then discusses, with a good set of illustrations, how satellites have helped scientists to learn more about the shape of the Earth. The nontechnical approach makes the article suitable for high school and general readers.

Melchior, Paul. *The Earth Tides*. Oxford, England: Pergamon Press, 1966. A sophisticated treatment of the physical phenomenon of small distortions of Earth resulting from gravitational forces imposed by the Moon and the Sun.

Stacey, Frank D. *Physics of Earth*. 2d ed. New York: John Wiley & Sons, 1977. A reference volume on solid-Earth geophysics, including radioactivity, rotation, gravity, seismicity, geothermics, magnetics, and tectonics. Provides detailed numerical tabulations on dimensions, properties, and unit conversions.

See also: Earth-Moon Relations; Earth-Sun Relations; Earth's Age; Earth's Core; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Rotation; Earth's Structure; Planetary Orbits: Couplings and Resonances; Planetary Rotation.

Earth's Structure

Category: Earth

Processes that occur in the interior of the Earth have profound effects upon the surface of the Earth and its human population. The results of such processes include earthquakes, volcanic activity, and the shielding of life-forms from solar radiation.

OVERVIEW

A simple demonstration that the Earth's interior is different from its surface is to compare the Earth's average density, calculated by dividing its mass by its volume, with the density of typical rocks from the surface. The average density is about 5.5 grams per cubic centimeter, while the density of typical surface rocks is between about 2.7 and 3.0 grams per cubic centimeter. This means that part of the interior must be composed of much denser material than surface rocks.

Evidence that the interior is differentiated into "layers" of various thicknesses, compositions, and mechanical properties comes primarily from analyzing the seismic waves produced by earthquakes that travel through the Earth. The thinnest layer is the outermost one known as the crust. The crust varies in thickness from about 5 kilometers under parts of the ocean basins up to about 70 kilometers under the highest mountain ranges of the continents. The crust is composed of a number of different rock types, but there are systematic differences between the crust of continents and that of ocean basins; continental crust is generally granitic (similar to granite), while oceanic crust is generally basaltic (similar to basalt). Both granite and basalt are igneous rocks, meaning that they cooled and hardened from hot molten material, and both are composed of silicate minerals.

However, the silicate minerals in basalt (such as pyroxene, olivine, and calcium-rich plagioclase feldspar) are comparatively rich in iron, magnesium, and calcium, giving them a generally dark color and slightly greater density (about 3.0 grams per cubic centimeter), while the silicate minerals in granite (such as quartz, potassium feldspar, and various micas) are poorer in iron, magnesium, and calcium, making them generally lighter in color and slightly lower in density (about 2.7 grams per cubic centimeter). Note that this distinction does not hold completely: Basalt and similar rocks can be found on continents, and sediments weathered from granite and similar rocks can be found in ocean basins.

The base of the crust is marked by a boundary known as the Mohorovičić discontinuity, or Moho. It represents a change in density of the rock above and below it. Rocks just below the Moho are slightly denser, about 3.3 grams per cubic centimeter, than either continental or oceanic crustal rocks. The rocks below the Moho probably are peridotite. Peridotite, composed mostly of olivine and pyroxene, is similar to basalt, but it is richer still in iron and magnesium. Peridotite is thought to represent the general composition of the layer underlying the crust, called the mantle. The mantle comprises the bulk of the Earth, representing about 80 percent by volume.

In the upper mantle, at depths starting about 100 kilometers beneath the surface and extending down to about 410 kilometers, is a zone of less rigid and more plastic, perhaps even partially melted, material called the asthenosphere. The crust and the part of the mantle above the asthenosphere, acting as a rigid unit, are known collectively as the lithosphere. The change to plastic behavior in the asthenosphere occurs because temperatures there are close to the melting point of peridotite. Although temperature continues to increase below the asthenosphere, the greater pressures at greater depths are high enough to keep the rock from melting.

The asthenosphere is thought to play an important role in movements of the lithosphere above. According to the theory of plate tectonics, the lithosphere is divided into a number of plates about 100 kilometers thick that are in constant motion at speeds of up to several centimeters per year, driven by hot convection currents of material moving slowly in the plastic asthenosphere. The hot material rises along divergent plate boundaries marked at the surface by the volcanic ridge-rift system that extends through the ocean basins around the globe. The slowly moving convection currents in the asthenosphere then move laterally away from the ridge-rifts, carrying the lithospheric plates above away from the ridge-rifts. As it moves laterally, the asthenosphere cools, becoming denser and sinking back downward. The sites where the convection currents sink are places where lithospheric plates with ocean crust on top dive into the mantle in a process called subduction. At these sites, marked at the surface by trenches in the ocean basin floor, crustal rocks may be carried into the upper mantle to depths as great as 700 kilometers. Below this level, the rock may simply be too dense for the lithospheric plates to penetrate.

There are two lower boundaries within the mantle. At 410 and 660 kilometers below the surface, abrupt increases in density occur. Although one might suspect a change in composition to account for the jump in density, laboratory studies of rocks under pressure suggest an alternative explanation. The primary mineral in peridotite is olivine. The pressures at 410 kilometers and again at 660 kilometers collapse the crystalline structure and produce denser minerals with the same iron and magnesium silicate composition. At the pressure existing at 410 kilometers, olivine converts to the denser mineral called spinel, and at the even higher pressure at 660 kilometers, both spinel and pyroxene collapse to yet a denser mineral known as perovskite. Thus the changes occurring in the mantle to produce the asthenosphere and the discontinuities at 410 and 660 kilometers are not changes in composition but instead changes in physical properties caused by temperature and pressure. The density increases from about 3.3 grams per cubic centimeter at the top of the mantle to about 5.6 grams per cubic centimeter at its base.

The next layer beneath the mantle is the outer core. This layer begins at a depth of about 2,900 kilometers beneath the surface and continues to a depth of 5,100 kilometers. There is a large density increase across the core-mantle boundary, from 5.6 grams per cubic centimeter at the base of the mantle to about 10 grams per cubic centimeter at the top of the core. Iron is the only reasonably abundant element that would have the required density at the tremendous pressure of millions of atmospheres at these depths. However, pure iron would give too high a density, so iron mixed with about 15 percent nickel, sulfur, silicon, and possibly oxygen and even hydrogen has been suggested. At the pressures and temperatures that must exist in the outer core, iron alloys would be in a molten state. Complex currents of metallic iron alloy, generated in the fluid outer core by convection and the Earth's rotation, give rise to the Earth's main magnetic field through a geodynamo process.

The core-mantle boundary represents a composition change from the silicate minerals of the lower mantle to the metals of the core. The boundary is a sharp one, but whether it is smooth and spherical in shape or irregular with "hills" or "peaks" on its surface is not known. There is some evidence from seismology that the lower mantle within 100 kilometers of the core boundary is a transition zone with a change of properties. It may consist of a mix of mantle and core material that is less rigid than the mantle rocks above it.

The innermost layer of the Earth's interior is the inner core. This region has a radius of about 1,300 kilometers where there is a boundary with the outer core. Increasing pressure at these depths requires that the iron of the inner core is solid. It is thought the solid inner core continues to grow in size as iron in the molten outer core crystallizes as Earth slowly cools. Because the solid inner core is separated from the mantle by the molten outer core, it can rotate independently. Seismic studies suggest that the inner core rotates slightly faster than the mantle and crust.

METHODS OF STUDY

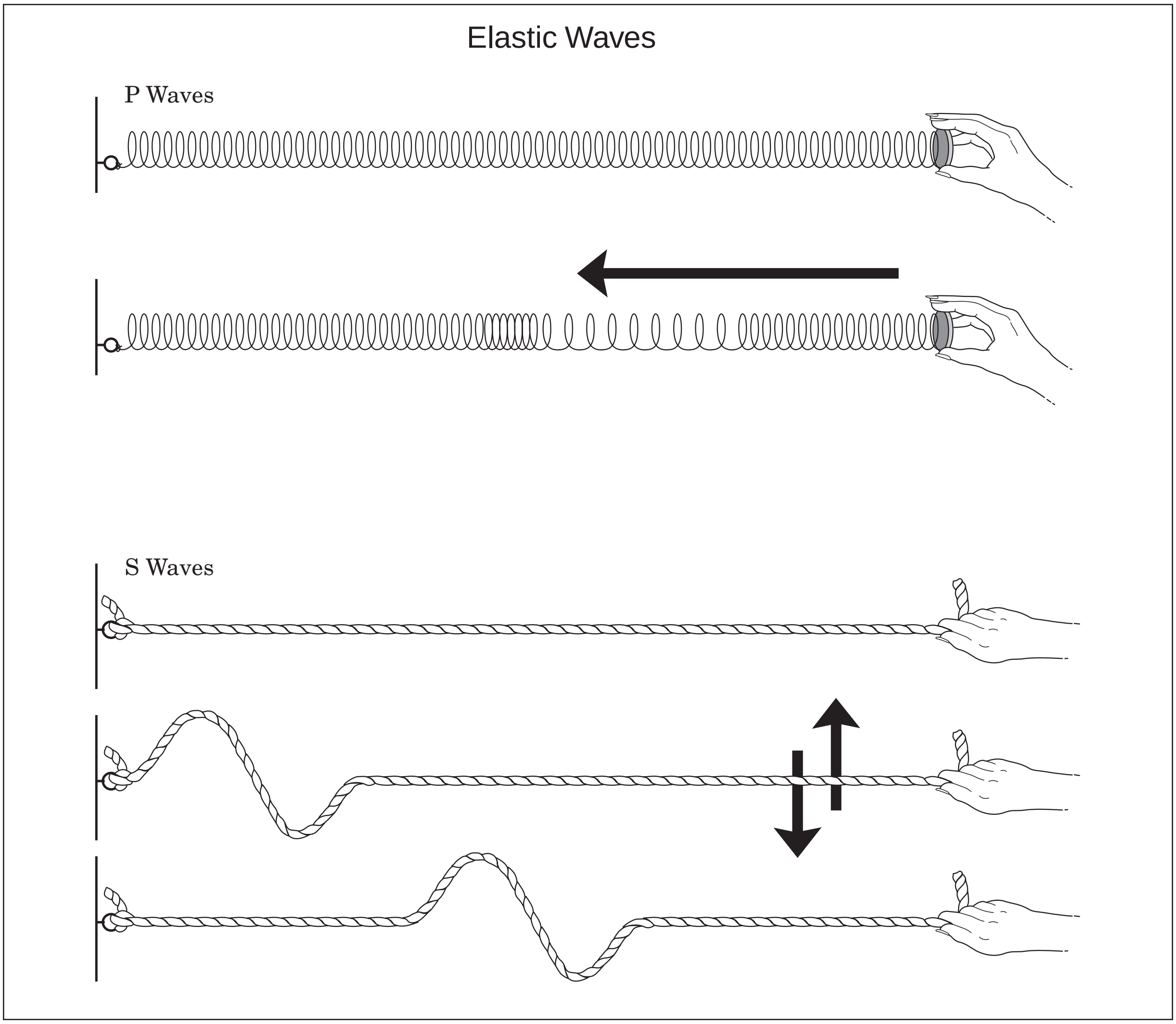
Much of what is known about the structure of the Earth's interior comes from the analysis of seismic waves generated by earthquakes or by explosives detonated at or just below the surface. After passing through the Earth, the wave vibrations are recorded on seismographs located all around the world, revealing information about the part of the interior they traveled through.

The seismic waves that pass through the interior are called body waves, because they propagate through the body of the Earth and not along the surface. Body waves are of two varieties: primary (or P) waves, and secondary (or S waves). P waves are the same as acoustic or sound waves. They cause the material they traverse to move back and forth in the direction of wave travel, alternately stretching and compressing it. Like ordinary sound waves, P waves can travel through any sort of material—solid, liquid, or gas. S waves are transverse waves, which move material along the wave path from side to side. Consequently they can travel only through rigid, that is, solid, material; S waves cannot travel through liquids or gases.

Both P waves and S waves cross the asthenosphere of the upper mantle, but with reduced speed (that is why the upper mantle is sometimes called the low-velocity zone), suggesting lower rigidity but not a liquid state, since S waves do propagate through it. Therefore, it seems that the asthenosphere is a solid but plastic region, able to ooze and flow very slowly.

At the core-mantle boundary, P waves abruptly slow down by almost a factor of 2 as they enter the outer core, and S waves disappear, indicating the material of the outer core has no rigidity. Since gases cannot exist at the conditions of the outer core, the material of this region must be a liquid.

Other locations in the interior are marked by increased speeds for both P waves and S waves.



There is a sharp increase in speed at the Mohorovičić discontinuity at the base of the crust. The P-wave speed jumps from about 6 kilometers per second above the Moho to around 8 kilometers per second below it, while the S-wave speed jumps from about 4 to 5 kilometers per second.

Below the asthenosphere, both P-wave and S-wave speeds gradually increase to a depth of about 410 kilometers, at which point both sharply increase. Laboratory studies indicate that the increase in density when olivine collapses to form spinel accounts for the increased speeds at this depth. At a depth of around 660 kilometers, a second abrupt increase in both wave speeds occurs, here caused by a second collapse to produce the yet denser mineral perovskite. The speeds for this part of the mantle match those obtained in the laboratories from waves passing through perovskite samples placed under the kinds of pressures found at 660 kilometers. A final increase in speed is observed when P waves pass the outer to inner core boundary. This is due to a phase transition from liquid iron in the outer core to solid iron in the inner core. Such a phase transition is supported by the probable reappearance of S waves in the inner core.

Another way in which the existence of structural boundaries within the Earth can be shown from an analysis of seismic waves is to examine the way the waves reflect and/or refract when they encounter the boundaries. What the waves do depends on the angle at which they approach the boundary, as well as on the properties of materials on both sides of the boundary. P waves are reflected off the Moho, the core-mantle boundary, and the inner-outer core boundary, providing clear evidence that there are sharp boundaries between the crust and mantle, the mantle and outer core, and the outer and inner cores. Waves have also been detected reflecting off the 660-kilometer discontinuity.

Refraction or bending of waves yields further evidence. As P waves cross the mantle-core boundary, they are refracted toward the center of the Earth because their speed is less in the molten outer core. This deflection of P waves passing through the outer core leaves a gap stretching around the Earth in the form of a band extending from 100° to 140° from the epicenter of the earthquake. This gap is known as the P-wave shadow zone because no P waves reach the surface in this band.

Advances in computer science have allowed the identification of even subtler details about the Earth's interior. Computerized tomography is a technique used in medicine, in which X rays from all directions are analyzed in a computer to give a three-dimensional picture of the internal organs of the human body. Seismic tomography is an analogous approach that uses seismic waves that travel from earthquakes to seismographs around the world to map the Earth's interior. This technique includes both P and S body waves traveling through the interior as well as surface waves traveling along the surface of the Earth. By looking at the travel times of the different waves, scientists are able to compare speeds along different paths. Such an approach has resulted in maps of slow and fast regions of the mantle that probably represent less rigid (warmer) and more rigid (cooler) regions.

CONTEXT

The interior of the Earth has profound effects on the surface. The interior acts as a complex heat engine that is the driving force behind plate tectonics, resulting in the formation and evolution of oceanic and continental crust. In the process, earthquakes and volcanic activity occur that create hazards for the human population on the Earth's surface. Complete acceptance of the plate tectonic theory could not have occurred without the discovery of the asthenosphere, which makes the movement of the lithospheric plates plausible.

It is fortunate that the Earth has a magnetic field. Without it, the age of discovery and exploration would not have been possible, for navigation by magnetic compasses allowed voyages across uncharted oceans. It is now theorized that the Earth's magnetic field is generated in the molten metallic outer core by the geodynamo process. An important effect of the core-generated magnetic field is the changes it has undergone through time. In particular, at rather irregular intervals of geologic time, the magnetic poles reverse polarity; during such reversals, the magnetic field decreases in strength. Since the magnetic field shields life-forms on the Earth's surface from charged particles emitted by the Sun, there is some concern that such field weakenings during polar reversals could result in more cancers and genetic mutations. Some scientists suspect that polar reversals might be at least partly responsible for some of the mass extinctions that have occurred in the geologic past, as well as periods of rapid evolution. Thus, the surface of the Earth as well as the life-forms on it depend upon and are strongly affected by processes occurring within the Earth's interior.

David S. Brumbaugh

FURTHER READING

Bolt, Bruce A. "Fine Structure of the Earth's Interior." In *Planet Earth*. San Francisco: W. H. Freeman, 1974. An extremely well illustrated review of how seismic waves have been used to discover and define the various layers of the Earth's interior. Written at a general-interest college level. Little background or expertise in mathematics is required.

_______. *Inside the Earth: Evidence from Earthquakes*. San Francisco: W. H. Freeman, 1982. This book is written for undergraduate college students in physics and the Earth sciences and for nonspecialists interested in a more detailed summary of knowledge of the Earth's interior. The text is relatively free of mathematics and is clearly and well illustrated. It offers a concise and readable treatment of the use of seismic waves to discover and interpret the Earth's interior.

Fowler, G. C. *The Inaccessible Earth: An Integrated View to Its Structure and Composition*. 2d ed. New York: Chapman and Hall, 1993. A commonly used introductory geophysics text. Readers should have knowledge of basic calculus. Includes a series of questions and problems.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. There is a very good chapter on the Earth's interior, as well as several chapters on volcanoes, earthquakes, and plate tectonics that describe their relationships with the interior structure and processes.

Van der Pluijm, Ben, and Stephen Marshak. *Earth's Structure*. 2d ed. New York: W. W. Norton, 2003. An introductory text on structural geology and tectonics. Designed for undergraduate students.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996. Covers the field of geophysics from 1960 with the theory of plate tectonics through the mid-1990's. Written with little technical jargon and many examples.

See also: Earth-Moon Relations; Earth-Sun Relations; Earth System Science; Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Rotation; Earth's Shape.

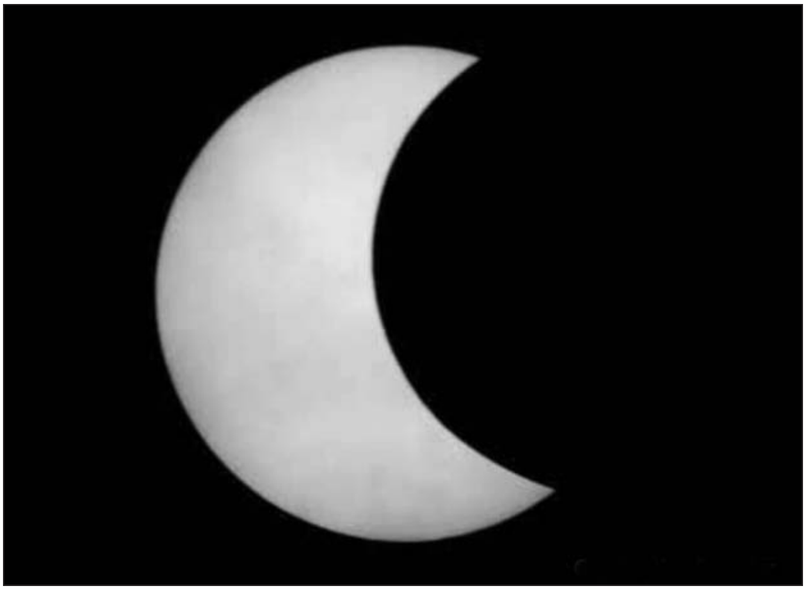
Eclipses

Categories: Earth; The Solar System as a Whole

Eclipses, occultations, and transits occur when three celestial bodies line up, causing the middle body to block the path of light between those on the two ends. In particular, solar and lunar eclipses witnessed from Earth are spectacular phenomena that have been objects of awe, study, and speculation since ancient times. Once they were understood, they became powerful tools of science used to investigate topics as diverse as geodesy and general relativity.

OVERVIEW

Eclipses of the Sun and Moon are impressive events. They have captivated people since before recorded history and continue to excite us today. Once considered great omens or portents, they have become among the most powerful means with which science tests theories in a remarkable variety of areas.



A partial solar eclipse was visible from Earth as the Moon obscured the Sun in 1994. (©Sébastien Gauthier/NASA)

A lunar eclipse (or eclipse of the Moon) occurs when the Moon passes through the shadow of the Earth. For this to happen, the Moon must be on the side of the Earth opposite the Sun, at the time when the Moon's phase is full. A solar eclipse (or eclipse of the Sun) occurs when the Moon passes between the Earth and the Sun, blocking all or part of the Sun as seen from Earth. For this to happen, the Moon's phase must be new.

Not every full or new Moon results in an eclipse, since the orbit of the Moon lies in a plane tilted about 5° to the ecliptic plane, the plane containing Earth's orbit around the Sun and in which the Earth and the Sun always lie. Unless a full or new Moon occurs when the Moon is very close to crossing the ecliptic plane, the Earth's shadow will miss the Moon (at full Moon) or the Moon's shadow will miss the Earth (at new Moon), and no eclipse will take place. This condition has been known and used to predict eclipses since ancient times and is the source of the name of the ecliptic plane.

A lunar eclipse is visible from all points on Earth where the Moon is above the horizon. It may be either umbral or penumbral. During an umbral lunar eclipse, at least part of the Moon passes through the Earth's umbra, the dark inner shadow in which the Earth blocks light coming from all parts of the Sun. If the entire Moon passes through the umbra, it is called a total lunar eclipse; if only part of the Moon passes through the umbra, it is called a partial lunar eclipse. During a penumbral lunar eclipse, the entire Moon misses the umbra and passes only through the Earth's penumbra, a region of partial shadow surrounding the umbra in which Earth cuts off light from some but not all parts of the Sun. An observer on the Moon during a penumbral lunar eclipse would see part of the Sun covered by the Earth and part of the Sun extending beyond the edge of the Earth.

The Moon is dimmed slightly while in the penumbra, but it does not darken appreciably unless it enters the umbra. When the Moon enters the umbra, the previously bright surface of the full Moon darkens to a much dimmer reddish glow, illuminated only by sunlight that has been refracted and scattered around the Earth by Earth's atmosphere. The brightness and color of this illumination can vary markedly from one umbral lunar eclipse to another (from orangish red to a dull reddish brown to a ghostly brownish gray), depending on the atmospheric conditions on Earth. Occasionally, some areas of the Moon will seem less illuminated than others.

A total lunar eclipse can last several hours. From the time the Moon begins to enter the umbra, it takes about an hour for the eclipse to become total. Totality can last for nearly two hours, with another hour required for the Moon to leave the umbra entirely. The limb (or edge) of the Moon nearest the observer's eastern horizon enters the Earth's shadow first, and at the end of the eclipse it is this limb that brightens first.

Solar eclipses are somewhat more complex. By coincidence, the Sun and the Moon both have nearly the same apparent size or angular diameter—about one-half degree of arc—as viewed from Earth. The Sun's actual diameter is about 400 times larger than the Moon's, but the Sun is also about 400 times farther away from the Earth than the Moon is. Because the orbit of the Moon around the Earth and the orbit of the Earth-Moon system around the Sun are both slightly elliptical, the apparent (angular) size of the Moon and Sun varies as seen from the Earth. On average, the angular diameter of the Moon, 0.518°, is slightly less than the angular diameter of the Sun, 0.533°, as seen from Earth. However, when the Moon is at perigee (its closest approach to Earth), its angular diameter increases to 0.548°. When the Earth is at perihelion (its closest approach to the Sun), the Sun's angular diameter increases to 0.542°. Thus, when the Moon is near perigee, its angular size always exceeds that of the Sun, and if a solar eclipse occurs then, the Moon can completely cover the Sun.

The Moon's umbra is conical in shape, with its base at the Moon and narrowing to its apex (or tip) as it nears Earth. If the Moon is near perigee, the apex of its umbra will fall inside the Earth, and observers in the region of the Earth's surface within the umbra will see the silhouette of the Moon completely cover the Sun's visible surface (its photosphere); this is called a total

The shadow of the Moon darkens and moves across Earth during a solar eclipse; those on Earth's surface in the center of the shadow see a total solar eclipse, while those at the perimeters see a partial solar eclipse. (Centre National d'Études Spatiales) solar eclipse. The region of the Earth's surface within the umbra at any moment is quite small, never more than a few hundred kilometers across. Due to the orbital motion of the Moon around the Earth and the Earth's rotation on its axis, its umbra sweeps a path (the eclipse track) thousands of miles long across the Earth's surface from west to east at speeds always exceeding 1,700 kilometers per hour relative to the Earth's surface. The maximum duration possible for totality is about seven and a half minutes, although the complete eclipse, including the partial phases at the beginning and end as the Moon slowly covers and then uncovers the Sun, may take more than four hours.

Because the Moon's average angular size is slightly smaller than the Sun's average angular size, the tip of the Moon's umbra will not always reach the Earth's surface during a solar eclipse. If this is the case, the Moon will not completely cover the Sun's photosphere. When the Moon is centered on the Sun, a narrow ring, or annulus, of the Sun's bright photosphere remains visible around the Moon's silhouette. This is called an "annular solar eclipse." Annular solar eclipses are about 20 percent more frequent than total solar eclipses.

Beyond the region of the Earth's surface in which a total or annular solar eclipse is seen, there is an area thousands of kilometers wide inside the Moon's penumbra. Within this area, the silhouette of the Moon covers part but not all of the Sun's photosphere. This is called a partial solar eclipse.

Occultations and transits are phenomena similar to eclipses in which the apparent angular size of the body in front is substantially larger or smaller than the apparent angular size of the body in back. An apparently large body moving in front of an apparently smaller one is called an occultation. The Moon frequently occults bright stars, which are seen to wink out instantly when they pass behind the limb of the Moon. The Moon also occasionally occults planets, and planets are seen to occult their satellites and stars. (Actually, other planets can both occult and eclipse their satellites as seen from Earth. The distinction between the two phenomena is that the satellite moving behind the planet is an occultation, while the satellite passing through the planet's shadow is an eclipse. The two events are not necessarily precisely coincident in time, because the planet's shadow cone will not be directly behind the planet as seen from Earth unless the planet is on the side of the Earth almost exactly opposite the Sun.) The Sun also occults objects, but because of the Sun's brightness, such solar occultations cannot be seen at visible wavelengths; however, they have been observed at radio wavelengths when the object being occulted is a source of radio emission. An apparently small body moving in front of an apparently larger one is called a transit. On rare occasions it is possible to witness from Earth the transit of the planets Venus or Mercury across the Sun. When this happens, the planet appears like a small, black dot moving across the face of the Sun. Transits of satellites and their shadows across their parent planets also can be observed.

Another, related phenomenon is that of eclipsing binary stars. Some of the stars seen in the sky are in fact pairs of stars orbiting about one another. If the Earth lies near the plane of their mutual orbit, the combined light of the system is seen to vary in brightness as the stars alternately block each other from Earth's view. By observing these variations in brightness, astronomers can determine some of the characteristics of the individual stars (such as relative sizes and surface temperatures) and study their interactions.

METHODS OF STUDY

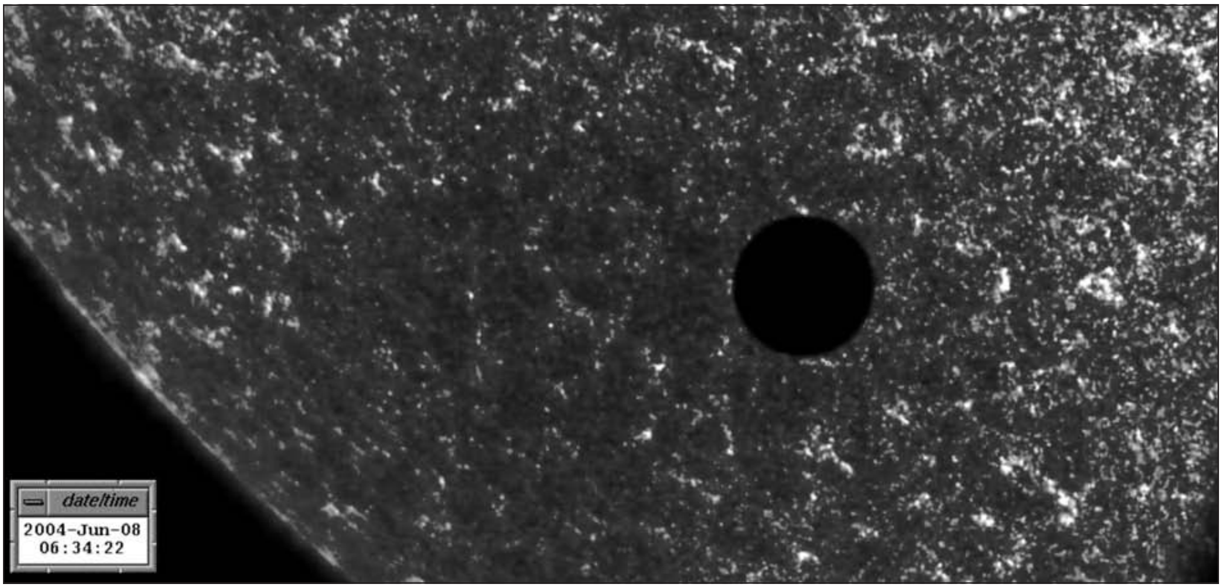
Ancient peoples, who used astronomical observations to keep track of planting seasons and the like, usually imputed magical or spiritual significance to eclipses and consequently tried to predict them. They did this by watching the changing position of the Moon against the background of the stars and by recording patterns in the recurrence of eclipses.

A lunar, or synodic, month is the period from one new Moon to the next, about 29.53 days. The draconic month is the time required for the Moon to complete one cycle of crossing and recrossing the ecliptic plane (from south to north and from north to south), about 27.21 days. The coincidence of these two cycles produces eclipses, so the pattern of eclipses starts again when the two cycles return to the same relative matchup. This happens every 223 lunar months (equaling 6,585.32 days or 242 draconic months or a little more than 18 years), in a repeating pattern called the saros cycle. The cycle lasts 18 years and 11.32 or 10.32 days, depending on whether 4 or 5 leap years occur during that time. The extra third of a day (or about 8 hours) means that eclipses 223 lunar months apart, although similar in overall geometry, will occur about 120° farther west in longitude because of the Earth's rotation. After three such saros cycles, about 56 years and 1 month, eclipses repeat in nearly the same part of the Earth again.

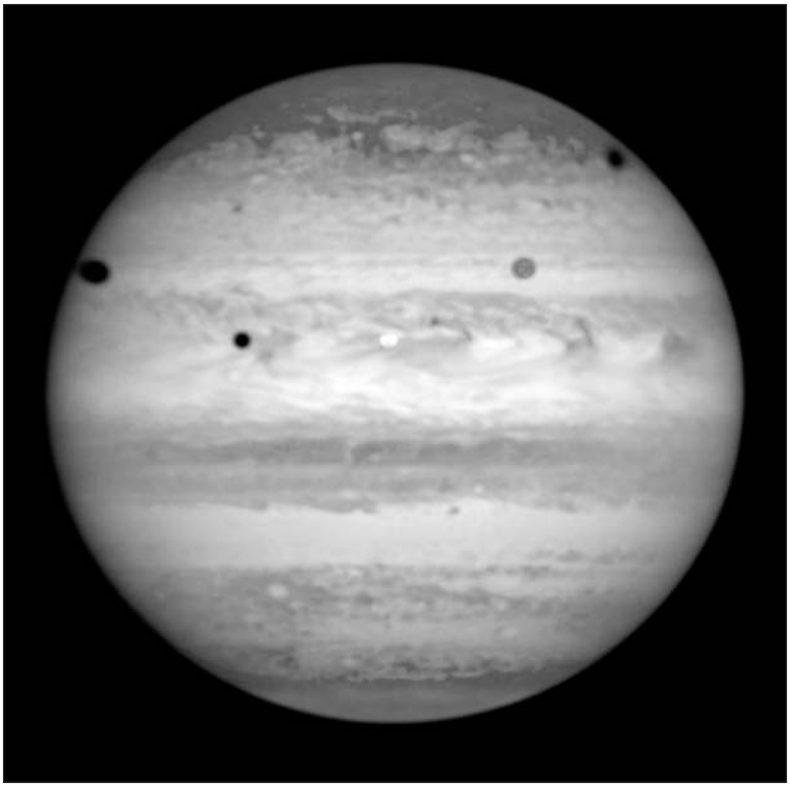
These cycles, at least as they applied to lunar eclipses, were known to Babylonian astronomers by around the eighth century B.C.E. and may have been known to some peoples long before that (based on disputed interpretations of a circle of fifty-six pits around the neolithic monument at Stonehenge in England). Knowledge of these cycles enabled the Babylonians to predict the relative motions of the Sun and Moon in the sky. The saros cycle was of limited use in predicting solar eclipses, because the path of totality is so narrow. Very precise knowledge of the relative motions of the bodies involved was required. This was not possible before Sir Isaac Newton developed his laws of motion and gravity along with the calculus in the seventeenth century. One of the first tests he applied to his new methods was the calculation of the orbit of the Moon.

Centuries of refinement, both of mathematical methods and of measurements of the positions of the Moon, Earth, and Sun relative to one another, were necessary to achieve modern accuracy in eclipse predictions. Astronomers can now calculate eclipses, including exact times and paths of totality, many years into the future with almost total precision. However, even these calculations are limited by residual uncertainties in the motions of the bodies involved when extrapolations hundreds of years in the past or future are attempted. With three bodies gravitationally interacting, no exact solution for the orbits is possible, although modern approximation methods are very good. Furthermore, the rate of rotation of the Earth has varied over time and continues gradually to slow, complicating the calculation of eclipse times and locations far back into the past or far forward into the future.

A total solar eclipse is perhaps the most spectacular natural event that can be seen. During a total solar eclipse, the Moon appears as a dark



The planet Venus can be seen crossing the solar face during its 2004 transit. (NASA) disk that slowly moves across and covers the bright disk of the Sun. Just before the Sun is completely covered, the remaining bright crescent narrows until it becomes a chain of bright spots along the edge of the Moon. These spots, called Bailys beads, represent a last glimpse of the Sun's photosphere between mountains at the edge of the Moon. Then for a few seconds, the Sun's chromosphere (a thin layer of transparent gas above the photosphere) can be seen as a red fringe along the leading edge of the Moon. At about this time, rapidly moving shadow bands, striations of light and dark a few centimeters across, can be seen rippling across the ground and along walls. These are believed to be due to atmospheric refraction. The sky turns dark during totality, but not completely dark; some light is scattered into the umbra from outside the region of totality. The darkness at totality produces uneasiness among some animals, and birds are sometimes seen to go to roost, as at sunset. The solar corona, the Sun's outer atmosphere of hot ionized gases, is seen as a glowing white halo around the dark silhouette of the Moon. Smaller, fiery red solar prominences are often observed around the edge of the Moon. As totality ends, the shadow bands, chromosphere, and Baily's beads can be seen briefly again.



The Hubble Space Telescope captures three of Jupiter's moons—Io, Ganymede, and Callisto—crossing the gas giant's face and casting shadows on its surface in a rare triple eclipse. (NASA/ESA/E.Karkoschka, University of Arizona)

It is important never to look directly or through optical instruments such as telescopes, binoculars, or camera viewfinders at the uneclipsed or partially eclipsed Sun without using suitable filters manufactured for this purpose. Common sunglasses are insufficient to prevent severe, painful, and permanent eye injuries, which can occur nearly instantaneously. However, during totality, when the Moon completely covers the Sun's photosphere and only the corona is visible, no filter is needed. Furthermore, no filters are needed to watch lunar eclipses. A telescope or at least binoculars generally are needed to enlarge the view of most occultations and transits.

CONTEXT

Eclipses, occultations, and transits have been powerful tools to derive useful information, to make new discoveries, and to test and confirm various theories and predictions. One of the oldest such uses can be traced back to the ancient Greeks. They understood how lunar and solar eclipses occurred, and the circular outline of the Earth's shadow seen on the Moon during lunar eclipses was cited by Aristotle (384-322 B.C.E.) and other Greek philosophers as evidence that the Earth must be spherical.

In 1675, the Danish astronomer Ole Rømer was studying the orbital motion of Jupiter's four largest moons (its Galilean satellites) by carefully timing their eclipses in Jupiter's shadow. He discovered that the eclipses occurred later than expected when Earth and Jupiter were farther apart and earlier than expected when Earth and Jupiter were closer. He realized that this phenomenon could be explained if light did not travel instantaneously but took time (about 16.6 minutes by his measurements) to cross the Earth's orbit (a distance of 2 astronomical units). Since the length of the astronomical unit was not known accurately then, he never actually calculated the speed of light in "everyday" type units, but this was the first demonstration that light traveled at a finite speed.

Eclipses of the satellites of Jupiter, as well as much less frequent lunar and solar eclipses, helped seafarers find their location at sea and map the Earth. Although latitudes can be determined easily by measurements of the maximum altitude above the horizon reached by the Sun in the daytime or specific stars at night, longitudes cannot be determined astronomically without a time reference. Pendulum clocks did not run accurately at sea because of the motion of the ship, but the calculated times of eclipses provided the necessary time reference and permitted reliable navigation and the construction of accurate maps.

Solar eclipses have helped resolve some of the most important questions in science. They provided an infrequent opportunity for observing some of the Sun's features—such as its corona, chromosphere, and prominences—which, before the development of modern instruments, were otherwise hidden most of the time by the brightness of the Sun. The element helium (from Helios, the Greek Sun god) was first discovered in the flash spectrum of the chromosphere during a total solar eclipse in 1868.

An observational test of general relativity was carried out first during the 1919 total solar eclipse. In 1915, Albert Einstein had published his then-controversial general theory of relativity. One of its predictions was that light passing by a massive object, such as the Sun, would be deflected by gravity. By photographing the star field around the totally eclipsed Sun and comparing the apparent positions of the stars near the edge of the Sun to a photograph of the same star field when the Sun was in a different part of the sky, astronomers verified this prediction of general relativity. It has been confirmed repeatedly at several total solar eclipses since then. Radio interferometry observations of a shift in the position of quasar 3C273 when it is occulted by the Sun have provided even more accurate confirmation.

In 1977, astronomers observed the occultation of a star by the planet Uranus to study its atmosphere by the way it absorbed light from the star. They were surprised when the star faded and brightened several times before and after being occulted by Uranus itself. This was attributed to a set of rings around Uranus, and was the first discovery of rings around a planet other than Saturn.

Firman D. King,
revised by Richard R. Erickson

FURTHER READING

Baker, Robert H. *Astronomy*. 7th ed. Princeton, N.J.: Van Nostrand, 1959. Baker's classic astronomy text provides a complete and lucid description of eclipse phenomena, along with related issues in spherical astronomy.

Brewer, Bryan. *Eclipse*. 2d ed. Seattle: Earth View, 1991. A good introduction to eclipse phenomena and their history. For the general reader.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A well-written college-level textbook for introductory astronomy courses. Has several pages on eclipses.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Includes a section on eclipses.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook, thorough and well written. A major part of one chapter deals with eclipses.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A thorough college textbook for introductory astronomy courses, divided into many short sections on specific topics. Eclipses are discussed in several sections.

Stevenson, F. Richard. "Historical Eclipses." *Scientific American* 247 (October, 1982): 170-183. This article discusses how historical records and astronomical calculations are compared to resolve questions in both history and astronomy.

Taff, Laurence G. *Celestial Mechanics*. New York: Wiley-Interscience, 1985. This book describes how orbits are calculated and makes it clear why the process becomes so complicated when long time intervals are involved.

Zirker, Jack B. *Total Eclipses of the Sun*. Expanded ed. Princeton, N.J.: Princeton University Press, 1995. Zirker explains how solar eclipses are observed and describes some of the scientific results that have been obtained from them.

See also: Auroras; Earth-Moon Relations; Earth-Sun Relations; Earth System Science; Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Core; Earth's Core-Mantle Boundary; Earth's Crust; Earth's Crust-Mantle Boundary: The Mohorovičić Discontinuity; Earth's Differentiation; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth's Mantle; Earth's Oceans; Earth's Origin; Earth's Rotation; Earth's Shape; Earth's Structure; Greenhouse Effect; Solar Chromosphere; Van Allen Radiation Belts.

Electromagnetic Radiation: Nonthermal Emissions

Category: The Cosmological Context

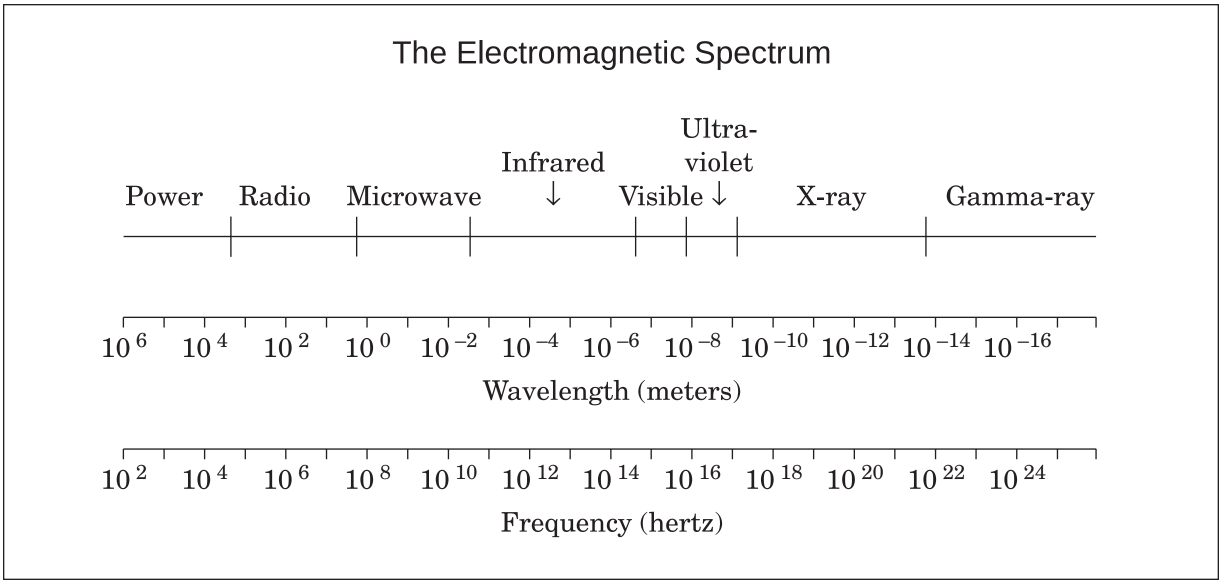
Objects emit light in many ways. Different conditions cause different forms of light to be emitted. Understanding the different forms of light emission can help in understanding the conditions where the light is emitted by studying the light itself.

OVERVIEW

Electromagnetic radiation is the name given to the phenomenon of waves of coupled electric and magnetic fields. Visible light is one form of electromagnetic radiation. Other forms, such as radio waves and X rays, are different solely in the frequency at which they oscillate and their wavelengths. The energy of electromagnetic radiation is linearly proportional to the frequency of the radiation. Different physical processes produce different forms of radiation. Among the most commonly produced electromagnetic radiation is thermal radiation (sometimes called blackbody radiation). Thermal radiation is produced by the molecular motion of anything that is not at absolute zero temperature. A plot of intensity versus wavelength for thermal radiation produces a well-defined characteristic spectral profile. However, there are many other physical processes that produce electromagnetic radiation other than through thermal means.

One of the most common forms of nonthermal electromagnetic radiation is emission spectra. When an electron in an atom jumps from a higher-energy state to a lower one, it emits electromagnetic radiation. The particular wavelength of radiation emitted depends on just the energy states in which the electron started and finished its transition. The greater the energy difference, the higher the frequency and the shorter the wavelength of radiation. Different atoms have very specific permitted energy levels for their electrons, and therefore they emit only specific radiative wavelengths. This is unlike thermal radiation, which is continuous across many wavelengths. All atoms of a particular chemical element have the same possible energy levels, so they will all have the same possible spectral line emissions, or spectral "signature" or profile. Thus, analysis of the spectral lines emitted by an object can be used to determine its chemical composition.

Another form of electromagnetic radiation emission that is similar to the emission spectra from electron transitions occurs through electron spin-flip transitions. Electrons have a small angular momentum, called the spin of the electron. This gives the electron a small magnetic moment. If that magnetic moment is



aligned with that of the atomic nucleus, then the atom has more energy than if it were aligned in the opposite direction. Therefore, if the electron flips its spin, the atom must emit a photon of light having energy equal to the difference in energy between the two states. One of the most important examples of this type of transition is the spin flip of hydrogen, which produces electromagnetic radiation with a characteristic wavelength of 21 centimeters.

Electromagnetic radiation comes from changing electric and magnetic fields, so any motion by a charged particle can produce electromagnetic radiation. In particular, certain types of electromagnetic radiation are produced from accelerated charges. One example of this type of radiation is synchrotron radiation. Acceleration is change in velocity. That change can be from speeding up or slowing down, or it can be from a change in direction of the velocity. When objects move around in circular or other curved paths, they change direction even if they do not change speed. If these objects are charged particles, then this curved motion results in the emission of electromagnetic radiation. This radiation is called curvature radiation or cyclotron radiation. However, if the charged particles are moving at or near the speed of light, the radiation is called synchrotron radiation and it takes on special properties. The synchrotron radiation is beamed along the path of the electrons, which amplifies its intensity, and it is polarized (meaning that the electric fields of the various electromagnetic waves are aligned). A spectrum of synchrotron radiation shows a drop in intensity with frequency, making an easily recognized spectral profile. A common source of synchrotron radiation is electrons or other charged particles moving in a magnetic field. Jupiter's magnetosphere emits synchrotron radiation.

Another form of radiation produced by accelerated electrical charges is Bremsstrahlung radiation. The name comes from the German meaning "braking radiation" because it is produced when fast-moving charged particles abruptly slow down. Technically, the term Bremsstrahlung radiation can be applied to any radiation emitted by an accelerating charged particle, including synchrotron radiation; however, in practice the term is frequently limited to cases where the particles slow down (negative acceleration). Bremsstrahlung radiation can be generated, for example, during solar flares. Energy released in the flare event can accelerate electrons near the surface of the Sun to almost the speed of light. The electrons then interact with the outer layers of the Sun to slow down, producing a burst of X rays from the Bremsstrahlung radiation of their deceleration.

Like synchrotron radiation, Bremsstrahlung radiation has a spectral profile that decreases in intensity with increasing radiation frequency.

Another form of nonthermal emission is Cherenkov radiation, which is also produced by ultrafast charged particles. Nothing can travel faster than light in a vacuum. However, light travels more slowly through a medium than through a vacuum. When a charged particle moving near the speed of light enters a medium in which the speed of light is slower than the particle, the changing electromagnetic fields produced by the particle's motion augment each other to produce electromagnetic radiation moving in the same general direction as the particle's motion. The particle slows as a result. While this sounds much like Bremsstrahlung radiation, the physical process is different. High-energy cosmic rays produce Cherenkov radiation when they enter Earth's atmosphere.

Yet another nonthermal source of radiation is light amplification by stimulated emission of radiation (laser). An excited atom can be stimulated to emit radiation by the passage of a photon of light having the same wavelength as the light that would be emitted if the atom's electron were spontaneously to go from a high energy level to a lower one. The stimulated radiation is in the same direction and in phase with the stimulating radiation, amplifying its intensity. The atmosphere of Mars is capable of lasing infrared light. Laser light is characterized by the light's being coherent; that is, all of the light is the same wavelength, moving parallel and in phase.

Pair annihilation can provide yet another form of nonthermal electromagnetic radiation. All particles of matter have a corresponding particle of antimatter. When a matter particle comes together with its antiparticle, the particle and the antiparticle annihilate each other. The combined mass of the particle and the antiparticle is converted into energy through Einstein's famous relationship $E = mc^2$. The energy is realized by a pair of photons, or particles of electromagnetic radiation, moving in opposite directions from the annihilation site. The energy of the particles, and hence their wavelength and frequency, is determined by the mass of the particle and the antiparticle.

KNOWLEDGE GAINED

Electromagnetic radiation is produced in many different ways. The type of radiation produced is determined by the physical process producing the radiation. Measurement of that radiation and its characteristics can then be used to understand the source of the radiation.

For astronomers, one of the most important forms of electromagnetic radiation is the emission from electron transitions between energy levels in an atom or molecule. In fact, the electromagnetic spectrum produced by a collection of atoms or molecules doing this is simply called its emission spectrum. The emission spectrum of a body can be used to determine its chemical composition. Furthermore, if the electrons are in a low energy level, light shining on the body can excite electrons to higher energy levels. Only light having energy equal to the energy difference between energy levels can be absorbed. Thus, only certain wavelengths of light can be absorbed. These wavelengths are the same as those given off in emission spectra. Therefore, studies of the light absorbed and not reflected can also be used to determine chemical composition of a body.

The 21-centimeter radiation can be used to detect the presence of hydrogen atoms in space. Synchrotron radiation can be used to probe the magnetospheres of planets. Bremsstrahlung radiation is important in studies of planetary atmospheres and the Sun. Cherenkov telescopes have been constructed on Earth to study the nature of cosmic rays. Studies of the lasing properties of planetary atmospheres also lead to better understandings of those atmospheres.

Though some is visible light, much of the nonthermal electromagnetic radiation is in forms other than as visual light. Thus, studies of nonthermal radiation use techniques from all areas of astronomy: optical, radio, X-ray, infrared, and ultraviolet astronomy. In fact, many forms of this radiation could not be studied at all until astronomers had developed the tools to study nonoptical wavelengths of light. Now that it is possible to study all of these wavelengths and the character of many forms of nonthermal radiation is understood, our understanding of the universe has greatly improved.

CONTEXT

Astronomers study objects that are very far away. Unlike most other scientists, astronomers do not have the luxury of being able to have the object of their study in a laboratory setting, but rather have to look at it from afar. This means primarily studying the electromagnetic radiation of the objects. While robotic spacecraft are capable of conducting studies in situ on the surface of a world, only a very few planets have been probed in this manner. Other spacecraft have flown by planets or even been placed into orbit around those planets. However, even most of these spacecraft have conducted their studies using electromagnetic radiation. Understanding the nature of this radiation and how it is created can lead to a far greater understanding of the bodies that are producing that radiation.

Every object that is not at absolute zero temperature emits thermal radiation, and study of the thermal radiation emitted by a body is very important to understanding it. However, far more information can be gleaned from studies of the nonthermal radiation emitted. The tools and techniques of doing so have become among the most important tools in the astronomer's arsenal for understanding celestial bodies.

Raymond D. Benge, Jr.

FURTHER READING

Asimov, Isaac. *Understanding Physics*. London: Allen & Unwin, 1967. A quite old but still very good book that introduces the basic concepts of electromagnetism for the nonscientist. Even though it is dated, it does an excellent job of covering the basic principles without mathematics. Widely available in libraries.

Caroll, Bradley W., and Dale A. Ostlie. *An Introduction to Modern Astrophysics*. 2d ed. San Francisco: Pearson Addison-Wesley, 2007. A textbook designed for undergraduate astronomy majors. The book gives an excellent explanation of the uses of the information gained by studies of thermal and nonthermal radiation. A knowledge of calculus is assumed.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. An excellent and thorough introductory college astronomy textbook. An entire chapter is spent on the nature of light (electromagnetic radiation), including emission and absorption spectra. Later chapters discuss synchrotron radiation.

Robinson, Keith. *Spectroscopy: The Key to the Stars—Reading the Lines in Stellar Spectra*. New York: Springer, 2007. Written for the serious layperson and amateur astronomer, this book does a good job of explaining spectroscopy. Prior knowledge of the physics involved is not necessary to follow the text.

Rybicki, George B., and Alan P. Lightman. *Radiative Processes in Astrophysics*. New York: Wiley, 1979. An advanced book aimed primarily at upper-division undergraduate students or first-year graduate students; however, anyone with a firm grasp of physics and mathematics can follow it. The book is a number of years old, but the basic processes that produce thermal and nonthermal radiation have been understood since before the book was written.

Verschuur, Gerrit L. *The Invisible Universe: The Story of Radio Astronomy*. 2d ed. New York: Springer, 2006. A well-illustrated work for the layperson, with many photographs. The book does not deal directly with a description of nonthermal electromagnetic radiation, but it does show how 21-centimeter and similar forms of nonthermal radiation are used in radio astronomy.

See also: Big Bang; Cosmic Rays; Cosmology; Electromagnetic Radiation: Thermal Emissions; General Relativity; Greenhouse Effect; Interstellar Clouds and the Interstellar Medium; Milky Way; Novae, Bursters, and X-Ray Sources; Optical Astronomy; Radio Astronomy; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Telescopes: Space-Based; Universe: Evolution; Universe: Expansion; Universe: Structure.

Electromagnetic Radiation: Thermal Emissions

Category: The Cosmological Context

Electromagnetic radiation emitted from the surface of an object provides valuable information regarding the physical state of the object. Whenever the temperature of the object is above absolute zero, the intensity of the radiation emitted and its distribution with wavelength—its thermal spectrum—largely depend only on the absolute temperature of the object. Thus, the temperature of an object can be determined from its thermal emissions.

OVERVIEW

An isolated atom or ion possesses a unique set of discrete, quantized energy levels. Transitions from a higher energy level to a lower one result in the emission of a photon of electromagnetic radiation whose wavelength depends on the difference in energy between the two levels. Specifically, $E_{photon} = E_{higher} - E_{lower} = hc/\lambda$, where h is Planck's constant, c is the speed of light in a vacuum, and λ is the wavelength of the photon emitted. The inverse process, involving the absorption of a photon of the same wavelength, will cause the atom or ion to transition from the lower energy state to the higher one. Thus, if an atom or ion is excited to one of its high energy states by a collision with some other atom or by absorbing radiation, cascading down to lower energy states produces emissions of photons of various specific wavelengths corresponding to the energy gaps between pair combinations of energy levels. This emission (or absorption) spectrum identifies the atom or ion involved in the process.

A discrete spectrum, however, can be quite complex, because it depends not only on the temperature and chemical composition of the object but also on its density, the local gravity field, the object's speed, and whether all parts are in thermal equilibrium. When atoms crowd together at high density, however—as in a solid, a liquid, or even a gas in the interior of a star—the energy levels of each atom or ion shift from their isolated configuration. The resulting smearing of energy levels relative to their isolated configuration permits a continuum of transitions where formerly only discrete transitions existed. If the material is dense enough, no restrictions exist as to which wavelengths can be emitted or absorbed. The electromagnetic radiation emitted in this case is called thermal radiation. In an idealized case where the composition and surface texture of the object can be neglected, the electromagnetic radiation emitted is also referred to as blackbody radiation. The shape and characteristics of the resulting spectrum depend only on the temperature of the object emitting the radiation.

Experimental investigations in the last quarter of the nineteenth century showed that the thermal spectrum of a hot object is characterized by relatively low-intensity emissions at short wavelengths, a rapid rise in intensity peaking at some intermediate wavelength, followed by a gradual diminishing in intensity at longer wavelengths. Furthermore, Wilhelm Wien, by applying the laws of thermodynamics to electromagnetic radiation, showed in 1893 that the wavelength at which the peak in the thermal spectrum appears depends on the temperature of the object according to the relation $\lambda_{peak} = b/T$, where T is the absolute temperature of the object and b is a constant whose value is 2.898×10^{-3} meter kelvin. (Max Planck later showed that b is itself constituted from other fundamental constants of nature.)

In 1879, Jośef Stefan used the experimental measurements of thermal spectra to deduce an expression giving the irradiance, the total energy radiated per unit time interval by a blackbody per unit area of the emitting surface, as a function of the object's temperature. In 1884, Ludwig Boltzmann provided the theoretical framework for the expression, which came to be known as the Stefan-Boltzmann law. This relation linking the irradiance with temperature, $R = \sigma T^4$, contains a new constant σ whose value is 5.67×10^{-8} joule/second/meter^2/kelvin^4 and indicates that the amount of energy radiated by a blackbody increases rapidly with temperature. Like the constant b, σ too was later shown to be composed of fundamental constants of nature.

By 1900 Max Planck had synthesized the

prior experimental and theoretical information with his quantum hypothesis to deduce the precise relation that describes the blackbody spectrum of an object. The intensity $I(\lambda,T)$ of light emitted at a wavelength λ by a blackbody at temperature T is given as

$$I(\lambda,T) = 2hc^2/[\lambda^5(e^{hc/\lambda kT} - 1)]$$

a relation known as Planck's law. The constant k is Boltzmann's constant. Though these relations strictly apply only to blackbodies, many objects, to varying degree, approximate blackbodies, especially if their temperatures are very high or very low.

Tracing an object's thermal spectrum involves measuring the intensity of light through different filters that transmit electromagnetic radiation in relatively small bands of wavelengths. Modern technologies permit the fabrication of filters and detectors for any region of the electromagnetic spectrum. Typically, however, measuring the intensity of thermal emissions in only three or four wavelength bands suffices to determine the shape of an object's thermal spectrum and therefore estimate its temperature. Because the Earth's atmosphere itself filters incoming electromagnetic radiation, intensity measurements in parts of the infrared and radio regions and nearly all measurements at wavelengths shorter than the violet are best made from satellites in space outside the atmosphere. The solar system contains objects that produce phenomena in a wide range of temperature.

The Sun, the hottest object in the solar system, generates enormous amounts of energy by means of nuclear reactions. Its hot, dense interior produces a spectrum which is nearly blackbody. Highly energetic phenomena in the Sun's photosphere, chromospheres, and corona superimpose nonthermal features on the blackbody spectrum, however. The planets, too, are sources of thermal emissions, not only because of their warmth in absorbing and then reemitting solar radiation, but also because of the residual radiating energy left over from their formation and internal heat generated by the decay of radioactive nuclei. Scattered throughout the solar system, dust fills much of the space between planets and is plentiful in its outer regions. Far from the Sun, the temperature of the dust falls to just a few tens of degrees above absolute zero, resulting in thermal emissions in the far-infrared region of the spectrum. Detecting thermal emissions from cold objects presents a challenge, however. In order to prevent the thermal emissions from the detector itself from overwhelming the signal from the object studied, the detector with its electronics must be cooled to temperatures very near absolute zero.

KNOWLEDGE GAINED

Thermal spectra provide a wealth of fundamental information regarding objects in the solar system. The Sun's spectrum is particularly informative. From satellite observations outside the complicating effects of Earth's atmosphere, the wavelength of the peak intensity is approximately 470 nanometers. Wien's law indicates that the corresponding temperature of the visible portion of the Sun, its photosphere, is 6,170 kelvins. Another estimate of the Sun's temperature may be obtained from the Stefan-Boltzmann law, which links the irradiance of the Sun to its temperature. From satellite measurements outside Earth's atmosphere, 1,368 watts of electromagnetic power are delivered to each square meter of detector surface. Because the distance to the Sun is known, the total power output of the Sun can be computed to be 3.85×10^{26} watts, and from its size, the irradiance is determined to be 6.34×10^7 watts/meter^2. The Stefan-Boltzmann law implies that the Sun's temperature is 5,780 kelvins. The temperatures determined by these two methods differ because the Sun is not a perfect blackbody. Nonthermal absorption features created by atoms and ions in its atmosphere remove energy from one part of the spectrum, and that energy reappears in other parts, thus producing a spectrum slightly distorted from that of a blackbody.

A combination of both Wien's law and the Stefan-Boltzmann law yields an equilibrium temperature profile within the solar system illuminated by the central Sun. Thus, objects farther from the Sun capture less of the Sun's radiant energy, are cooler than objects nearer the

Sun, and so possess thermal spectra whose peak wavelengths are longer than objects nearer the Sun. As an example of an application of this principle, the location of the majority of the dust in the solar system can be deduced. Visual observations from Earth have long recognized that the zodiacal light was likely due to dust lying in the plane of the ecliptic. Infrared satellite observations of the dust indicate that the peak in the thermal spectrum of the dust occurs at approximately 12 microns (1 micron = 10^{-6} meters). The 12-micron peak in the thermal spectrum of the dust in the solar system corresponds to a distance of 4.0×10^{11} meters, or about 2.7 AU (one astronomical unit, or AU, being the mean distance between Earth and the Sun), a location at the inner region of the asteroid belt. Most likely, collisions among objects within the asteroid belt generate the observed dust.

CONTEXT

An idealized blackbody spectrum quite closely mimics the spectrum of thermal electromagnetic radiation produced by a real object for the corresponding temperature. In this case, the wealth of theoretical results associated with blackbody radiation can be used to tease out information from the object under study. Within our solar system, images of rocky worlds taken through two or three filters in different parts of their thermal spectra produce a composite image that displays the distribution of minerals on their surfaces. In some instances, such a composite image identifies the minerals without having to retrieve a sample. This helps planetary scientists understand the relation between the mineralogy and the landforms, providing valuable clues as to the processes that produced the visible landscape. In other cases, thermal emission images of the surface of a planet, satellite, or asteroid serve as an important prelude to identifying the best target areas for a landing and site exploration, as in the case of many missions to the Moon and Mars.

The same information gleaned from the thermal spectra of objects within the solar system can be retrieved from a study of the spectra of objects and systems throughout the cosmos. Dust around young stars sometimes points to the formation of new solar systems beyond our own. Finally, the universe itself announces itself in subtle, omnipresent thermal emissions corresponding to a blackbody at a temperature of 2.725 kelvins, the radiation left over from the hot, dense creation of the cosmos 13.7 billion years ago.

Anthony J. Nicastro

FURTHER READING

Beiser, Arthur. *Concepts of Modern Physics*. 6th ed. New York: McGraw-Hill, 2002. An excellent text that presents a conceptual, accessible introduction to modern physics. Chapter 2 presents a unified overview of electromagnetic radiation and the development of theories to understand this phenomenon. Stresses the atomic nature of processes.

Bennett, Jeffrey, Megan Donahue, Nicholas Schneider, and Mark Voit. *The Cosmic Perspective*. 4th ed. San Francisco: Pearson/Addison-Wesley, 2007. A brilliantly illustrated, coherent, well-written introduction to astronomy. Chapter 5 contains a general overview of light and its interactions with matter. Thermal spectra, along with other types of spectra, are linked. Other parts of the book show applications to the solar system and the cosmos.

Cole, George H. A., and Michael M. Woolfson, eds. *Planetary Science: The Science of Planets Around Stars*. Bristol: Institute of Physics Publishing, 2002. Discusses an extensive array of applications of thermal and nonthermal spectra to objects in our solar system and in extrasolar systems. Primarily a book that covers descriptive planetary science, requiring very little prior knowledge of physics, astronomy, or mathematics.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Fransisco: Pearson/Addison-Wesley, 2003. With an early introduction to thermal radiation laws, this volume applies them to the study of all the objects in the solar system. Primarily descriptive and conceptual, the book makes occasional use of algebra.

Van der Meer, Freek D., and Steven M. De Jong, eds. *Imaging Spectrometry*. New York: Kluwer Academic, 2002. Provides a comprehensive description of the principles and applications of satellite and airborne spectrometry. Gives an extensive account of data collection and interpretation of images of geological and agricultural regions. Focuses primarily on terrestrial applications.

Zeilik, Michael, and Stephen A. Gregory. *Introductory Astronomy and Astrophysics*. 4th ed. Fort Worth, Tex.: Saunders College Publishing, 1998. A college-level text that provides broad coverage of astronomy and astrophysics. Results from an early cursory coverage of thermal radiation are used throughout an examination of the solar system. Later in-depth coverage of electromagnetic thermal and nonthermal radiation forms the basis for studying objects and phenomena across the cosmos. Makes heavy use of algebra and trigonometry.

See also: Big Bang; Cosmic Rays; Cosmology; Electromagnetic Radiation: Nonthermal Emissions; General Relativity; Greenhouse Effect; Interstellar Clouds and the Interstellar Medium; Milky Way; Novae, Bursters, and X-Ray Sources; Optical Astronomy; Radio Astronomy; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Telescopes: Space-Based; Universe: Evolution; Universe: Expansion; Universe: Structure.

Enceladus

Categories: Natural Planetary Satellites; The Saturnian System

Enceladus is the brightest of the satellites located within the rings of Saturn. It has much in common with what scientists expected to find on comets, especially water ice. However, it appears to have been formed more than four billion years ago as a spinning mass of soft material.

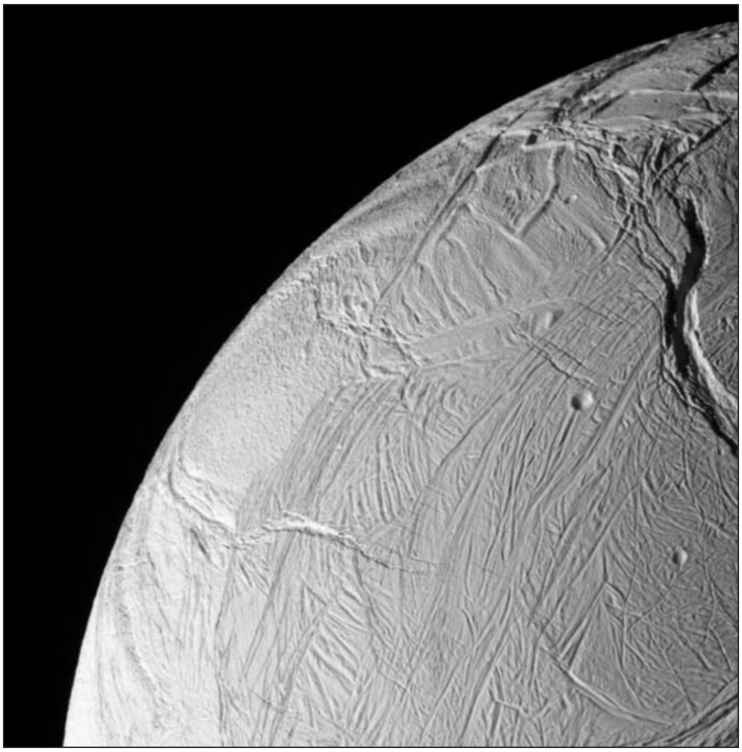
OVERVIEW

Enceladus was discovered on August 28, 1789, by William Herschel. Its orbit around Saturn has a semimajor axis of 237,948 kilometers and an eccentricity of 0.0047, with a period of 118,386.82 seconds (nearly 33 hours). The orbit is inclined at 0.019° to Saturn's equator and located inside the E ring around Saturn. This places Enceladus at roughly 4 Saturn radii, located between the orbits of the moons Mimas and Tethys, which are about one Saturn radius on either side. Its rotation is synchronous and has no axis tilt, so that the same hemisphere always faces Saturn. It is nearly spherical, with a mean diameter of 504.2 kilometers, being a slightly flattened ellipsoid with dimensions of 513.2 kilometers along the orbit radius pointed at Saturn, 502.8 kilometers along the orbit path, and 496.6 kilometers between the north and south poles.

It is the sixth-largest moon of Saturn. It has a mass of 1.08022×10^{20} kilograms, and the mean density is approximately 1,609 kilograms per cubic meter. The value of acceleration due to gravity at the surface on the equator is 0.111 meter/second^2. The escape velocity at the surface is 0.238 kilometer/second, neglecting atmospheric drag. Enceladus has very high Bond albedo of 0.99 and geometric albedo of 1.375, the highest among the satellites embedded in the Saturnian rings, indicating strong reflection. Its apparent magnitude from Earth is 11.7.

Enceladus is tidally locked in synchronous orbit around Saturn, meaning that the hemisphere with the higher density always faces Saturn. Looking into the sky from the Saturn-facing side, the planet would occupy roughly 30° of the sky and appear to be spinning in roughly the same position at all times. From the side hidden from Saturn, the Sun would appear very small, rising and setting in roughly 17 hours as Enceladus completed half an orbit around Saturn.

The orbital eccentricity of Enceladus is attributed to a resonance with the satellite Dione, with Enceladus completing two orbits for each orbit by Dione. This resonance may also drive the tidal heating of Enceladus. The shape of Enceladus is very close to that of an equilibrium-flattened ellipsoid, hydrostatically balanced by gravity and spin, which is the shape that an object would have in space if it were composed of homogeneous and fluid material. However, simulations of orbital evolution and the tidal locking suggest some variation in its internal density. The higher average density than that of water indicates denser material, possibly silicates, inside Enceladus.



The smooth, young surface of Enceladus is seen in this October, 2008, image from the Cassini spacecraft. (NASA/JPL/Space Science Institute)

Most of the data about the surface and orbital environment of Enceladus come from close flyby observations by the Voyager and Cassini spacecraft. The Voyager mission revealed evidence of a complex thermal history of Enceladus and showed several provinces with distinct geographical features. Short periods of intense heating and geological activity appear to have been separated by long periods of inactivity. Surface features include long, narrow depressions (fossae), ridges with cliffs of several hundred meters (dorsa), plains (plantia), long, parallel canyons (sulci), and craters. The terrain of the northern latitudes appears to be more than 4.2 billion years old, with more than one crater every 5 square kilometers, most of them bowl-shaped. Some craters are as large as 35 kilometers in diameter, while the majority are very small, less than a meter in diameter. The equatorial plains, named the Sarandib Plantia, show striations and folding, with about one crater every 70 square kilometers. Craters in these younger regions show viscous relaxation, indicating mechanisms for melting or distortion of the surface.

The ridged and grooved plains of the Samarkand Sulcus, at 55° to 65° south latitude, are 100,000 to 500,000 years old, and the fractured regions south of that show few if any craters. The Cassini orbiter approached the south polar region of Enceladus to within 168 kilometers on July 14, 2005, taking images with a resolution of 4 meters per pixel. House-sized boulders believed to be made of ice littered the polar landscape, but craters were nearly absent. A set of parallel "tiger-stripe" fractures, roughly 500 meters deep, 2 kilometers wide, and 130 kilometers long, are flanked by 100-meter high ridges. Temperature is as high as 175 kelvins. South of 55° south latitude, a chain of fractures and ridges circumscribes the moon. Some fractures intersect and overlay others. Scientists associate these fractures and ridges with the flattening and extension due to gravitational interactions with Saturn and other moons.

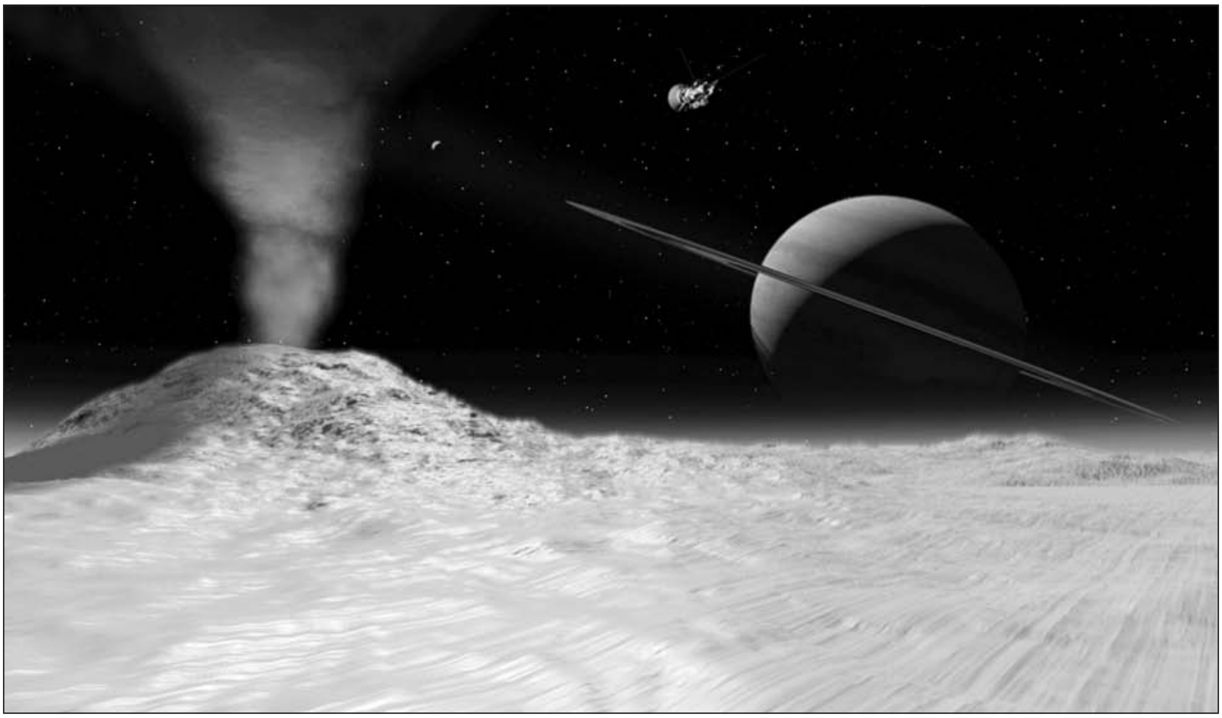
The Cassini flyby of Enceladus in 2005 revealed a water-rich plume ejecting as narrow jets from vents in the tiger-stripe sulci of the south polar region. The fine sizes of the particles in the plume, which freeze to ice or sublimate soon after ejection, suggest that the plume originates in a subsurface body of liquid water. Some of the water jets reach exit speeds of 600 meters per second, well above the 238 meters per second needed to escape the gravity of Enceladus. The mass flow rate of the plume is on the order of 100 kilograms per second, comparable to that of air through a modern supersonic jet fighter engine. Much of this mass may initially escape Enceladus into the ring around Saturn and may be the origin of the mass in the E ring of Saturn.

Enceladus also captures mass as snow falling to the surface from the E ring. Some scientists think that less than 1 percent of the mass in the plume can escape into the E ring. Numbers cited in different papers vary greatly on this subject.

Cassini approached to within 52 kilometers of the southern middle latitudes on March 12, 2008, in its E3 close approach, then dropped behind Enceladus in its orbit. It grazed the edge of the plume, which trails the satellite like the tail of a comet. The density of these portions of the plume had been predetermined from Earth by observing the dimming of starlight as the plume crossed in front of a star with an ultraviolet imaging spectrograph (UVIS). Roughly 70 seconds after closest approach, the craft was 250 kilometers above the surface, when the onboard ion and neutral mass spectrometer (INMS) encountered a peak particle density of nearly 10 million particles per cubic centimeter. This was still in the outer edge of the plume. The sharply defined plume edge provides further strong evidence that the plume comes out as a supersonic jet, as opposed to a diffuse subsonic plume caused by friction heating of ice due to the tidal stresses at the fractures. Measured gas density was twenty times higher than that predicted based on thermal expansion.

KNOWLEDGE GAINED

Data from the INMS suggest that the plume contains, beyond nearly pure water (ice), significant amounts (ranging from 1 to 10 percent) of carbon dioxide, methane, and other organic molecules, both simple and complex. There was a strong signal from something of molecular weight near 28, but there is debate over whether this substance is nitrogen (per the INMS data) or carbon monoxide (per the Cassini plasma spectrometer data). The values were similar to those predicted for many comet tails. The organic molecules detected include acetylene (C_2H_2), ethane (C_2H_6), hydrogen cyanide HCN, formaldehyde (H_2CO), propyne (C_3H_4), propane (C_3H_8), and acetonitrile (C_2H_3N). The cosmic dust analyzer instrument on the craft did not succeed in capturing the particle sizes. The Cassini plasma spectrometer (CAPS),



An artist's conception of the surface of Enceladus, showing one of the moon's ice geysers and Saturn in the background. (NASA) which measures ions, detected much larger particles as well, on the order of nanograms. If these were ice particles, they may have been as large as 0.01 to 0.1 millimeter in diameter, comparable to the particle sizes calculated from surface observation data of the south polar terrain. CAPS also detected positively and negatively charged ions, segregated in different regions of the plume.

Enceladus has no measurable internal magnetic field, but it has a significant influence on the magnetosphere of Saturn, strongly deflecting magnetic lines. This is now attributed to the water plume. Ions accelerated to energy levels of 20 kilo-electron volts by Saturn's strong magnetic field collide with the molecules in the plume, breaking them up into atoms and ionizing them. These fresh ions are again accelerated by the magnetic field, and, in turn, Enceladus substantially deflects the magnetic field lines of Saturn. Scientists also associate clouds of oxygen and hydrogen observed around Saturn with the atoms and ions generated when the water molecules in the plume from Enceladus collide with high-energy ions in the E ring.

Enceladus shows a trace atmosphere with surface pressure varying significantly in spatial location and perhaps in time, composed of about 90 percent water vapor, 4 percent nitrogen, 3.2 percent carbon dioxide, and 1.7 percent methane. Surface temperature varies from 32.9 to 145 kelvins, with a mean of 75 kelvins. The surface appears to be covered in clean water ice, accounting for the high reflectivity. The atmosphere is hypothesized to be an expanding, supersonic neutral gas cloud emanating from the surface after molecules are "sputtered" from the surface by collisions with high-energy ions.

In late 2008, presentations given at the American Geophysical Union meeting in San Francisco revealed that data from Cassini's observations of Encedalus strongly indicated that the satellite's surface displays action similar to the action of Earth's ocean floor, where new crustal material emerges from slits in the crust. Cassini imaging team leader Carolyn Porco proposed that liquid water was present on Enceladus's surface and that that surface splits and spreads apart. On Earth, molten rock rising up from deep in the planet causes the crustal

Enceladus's "tiger stripes" are visible in this Cassini image from 2008. (NASA/JPL)

spreading, whereas on Enceladus the surface spreading originates with upwelling of liquid, presumably water. Evidence suggested that the Tiger Stripes formations near the satellite's south pole are akin to the mid-ocean ridges found on the Earth's seafloor. Close flybys of that region resulted in more data on eruptions of water through vents in the Tiger Stripes formation. Combined with evidence of crustal spreading, these data have revealed Enceladus to be a surprisingly active world.

CONTEXT

The presence of water ice in a low-gravity body within seven years' travel time of Earth excited planners of deep space missions, since water is an excellent future propellant. The discovery of high-speed water jets from the south polar region provides strong evidence of liquid water below the surface, and continuing tectonic processes. Complex organic molecules discovered in the jet plume fuel speculation about precursors of life. Enceladus is one of three known planetary bodies in the solar system (besides Earth and the Jovian satellite Io) that has an internal heat made visible by remote sensing. The relations between Enceladus and the E ring, the magnetosphere, and clouds of oxygen and hydrogen around Saturn are subjects of intense study.

The similarity between Enceladus and comets has raised questions about the origin and evolution of the solar system. Comets were thought to have originated far outside the orbit of Pluto, independently of the planets, while the planetary satellites formed from the same cloud as the Sun and planets. However, Enceladus appears to have nearly pure water ice, as predicted for comets. How such a large, nearly spherical cometary body could have been captured in an orbit so close to Saturn is an unanswered question.

Narayanan M. Komerath and David G. Fisher

FURTHER READING

Benna, M., and W. Kasprzak. "Modeling of the Interaction of Enceladus with the Magnetosphere of Saturn." *Lunar and Planetary Science* 38 (2007). Discusses results of different numerical models of the magnetosphere interaction and the atmosphere of Enceladus, comparing against the results from the Cassini instruments.

Khurana, K. K., M. K. Dougherty, C. T. Russell, and J. S. Leisner. "Mass Loading of Saturn's Magnetosphere Near Enceladus." *Journal of Geophysical Research* 112, A08203, doi:10.1029/2006JA012110, 2007. Reports on modeling of magnetic field data on the interaction between Saturn's magnetosphere and Enceladus. Gives results on mass pickup and current generated.

Porco, C. C., et al. "Cassini Encounters Enceladus: Background and the Discovery of a South Polar Hot Spot." *Science* 311 (March 10, 2006): 1401-1405. Discusses the initial discovery of the relatively warm regions near the south pole of Enceladus.

Verbiscer, A., R. French, M. Showalter, and Paul Helfenstein. "Enceladus: Cosmic Graffiti Artist Caught in the Act." *Science* 315 (February 9, 2007). Discusses the albedo of Enceladus compared to the albedos of other moons of Saturn.

Wilson, D., et al. "Cassini Observes the Active South Pole of Enceladus." *Science* 311 (March 10, 2006): 1393-1401. Presents images and discusses the reasoning regarding the features around the south pole of Enceladus, based on the Cassini spacecraft's close approaches in 2005. Also discusses how shape and size considerations are used to form hypotheses on the evolution of Enceladus.

See also: Callisto; Eris and Dysnomia; Europa; Ganymede; Iapetus; Io; Jupiter's Satellites; Planetary Satellites; Saturn's Satellites; Titan; Triton.

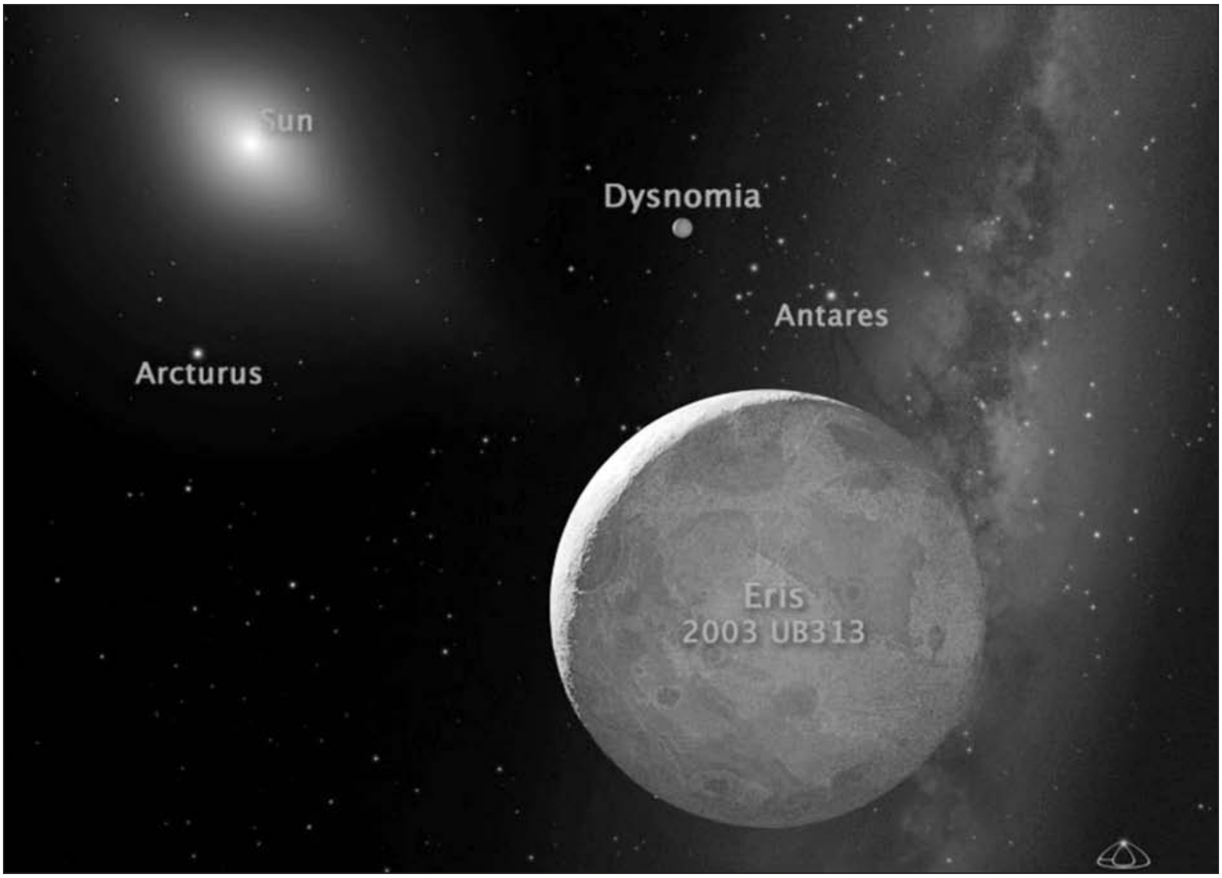
Eris and Dysnomia

Category: Small Bodies

The scattered disk object Eris is the largest dwarf planet and the most distant solar-system object astronomers have identified. Observation of Eris so far has revealed only one satellite, Dysnomia. Eris was initially hailed as a possible tenth planet when discovered in 2005, but debate within the astronomical community regarding what constitutes a planet led the International Astronomical Union to categorize it as a dwarf planet.

OVERVIEW

Eris, formally designated (136199) Eris, is the largest of the known trans-Neptunian objects. As of 2009, only one satellite, Dysnomia, had been found in orbit around Eris. The most distant solar-system object observed to date, Eris is classified as a scattered disk object. Despite its remoteness, it can be viewed with powerful amateur equipment. At aphelion, Eris lies beyond the outermost region of the Kuiper Belt; at perihelion, it passes within the range of Neptune's influence. As of 2009, Eris was one of only three known plutoids (the others are Pluto and Makemake) and four known dwarf planets (the three plutoids plus Ceres). Eris is larger than



An artist's view of dwarf planet Eris and its satellite Dysnomia, based on data from the Hubble Space Telescope. (NASA/ESA/Adolph Schaller for STScI)

Pluto, its orbital period is more than twice Pluto's, and at aphelion it is roughly three times farther from the Sun than Pluto.

Researchers have calculated Eris's orbit using archival data collected before the dwarf planet's discovery. Once Eris was identified in 2005, the bright, slow-moving object was easy to spot in images going back to 1989. Eris takes 557 years to travel around the Sun, following a highly elliptical path with a semimajor axis, or mean orbital radius, of 67.9 astronomical units (AU), or more than 10 billion kilometers. When Eris was discovered, it was near its aphelion of 97.5 AU (more than 14.5 billion kilometers) from the Sun. It will reach perihelion at 38.2 AU (more than 5.7 billion kilometers) in the mid-twenty-third century. For the most part, the orbit of Eris is typical of a scattered disk object. However, Eris's orbital inclination is an atypically high 44° from the solar system's orbital plane. The inclination and period of Eris's rotation have yet to be ascertained, although it is currently thought that a day on Eris is about eight hours long.

Images from the Hubble Space Telescope indicate that Eris has a diameter of roughly 2,400 kilometers, making it slightly larger than Pluto. Spitzer Space Telescope observations suggest that Eris's diameter may be as great as 2,600 kilometers. Calculations based on Keck Observatory and Hubble Space Telescope observations of Dysnomia's orbit around the dwarf planet indicate that Eris has a mass of 1.66×10^{22} kilograms, or 3.66×10^{22} pounds (27 percent greater than Pluto's) and a bulk density of 2.3 grams per cubic centimeter (0.83 pound per cubic inch).

Using Eris's density, researchers have determined that the dwarf planet's interior is most likely similar to Pluto's—that is, about half rock and half ice. Where Pluto's surface appears reddish and partly rocky, however, Eris has a uniform, highly reflective, almost gray surface. Near-infrared images obtained at the Gemini North Observatory in Hawaii indicate that this surface is predominantly frozen methane. In Pluto's case, darker surface hues are attributed to tholins, reddish-brown breakdown products formed when methane and similar organic compounds are subjected to solar ultraviolet irradiation. Where tholin deposits make Pluto's surface darker, the albedo is lower and the temperature higher. Methane ice melts away from these comparatively warm patches. By contrast, methane ice appears to envelop Eris in a bright, near-uniform coating. This suggests not only that Eris remains cold enough that its methane stays in a frozen state but also that there may be a subsurface source of methane that replenishes Eris's surface coating of methane ice and covers up whatever tholins are deposited.

As Eris moves from aphelion to perihelion, its temperature increases from its current value of 30 kelvins (–243° Celsius) to 56 kelvins (–217° Celsius). It is possible that, when Eris approaches the Sun, some of its surface ices become warm enough to sublimate and form a thin atmosphere. Whatever gases do not escape the atmosphere freeze once again as the dwarf planet moves toward aphelion.

To date, only one satellite has been observed in orbit around Eris: Dysnomia, known technically as (136199) Eris I. Hubble Space Telescope and Keck Observatory data indicate that Dysnomia is roughly 150 kilometers in diameter. It takes about 16 days for Dysnomia to complete its near-circular orbit around Eris at a distance of approximately 37,000 kilometers. The small satellite is believed to be composed largely of frozen water.

Astronomers Mike E. Brown of the California Institute of Technology, Chad A. Trujillo of Gemini Observatory, and David Rabinowitz of Yale University discovered Eris through an ongoing survey conducted at Palomar Observatory in southern California using the Samuel Oschin telescope. Images taken on the night of October 21, 2003, showed the large, bright object traveling slowly across the sky. Its movement was slow enough, in fact, that Eris went undetected when the images were first analyzed. The researchers' discovery in November, 2003, of Sedna, another large and slow-moving trans-Neptunian object, led them to adjust their detection scheme and reanalyze their survey data. On January 5, 2005, they identified the planet-sized scattered disk object that would later be known as Eris, designating it 2003 UB313.

Brown and his team intended to follow standard scientific protocols by verifying their discovery, studying it, documenting it thoroughly, and making it known through a scientific paper published in a reputable journal. However, in July, 2005, they learned that detailed records of their telescope use had inadvertently been made accessible to anyone with Internet access, and that an abstract they had recently published unwittingly contained clues about where in the sky to look for their recent—and unannounced—trans-Neptunian discoveries. When, five days after the abstract was issued, researchers in Spain announced the discovery of 2003 E161—a trans-Neptunian object Brown and his colleagues had also found—it appeared the team had to lay claim to 2003 UB313 or risk having someone else take credit for finding it.

The California Institute of Technology, the Jet Propulsion Laboratory, and the National Aeronautics and Space Administration (NASA) announced the discovery of 2003 UB313 on July 29, 2005, in press releases that referred to the object as the tenth planet. The media was quick to adopt the team's nickname for the newly discovered member of the solar system: Xena, so called for the heroine of the television series *Xena: Warrior Princess*. (Some overenthusiastic journalists, seeing "planetlila" in Brown's Web address, pounced upon Lila as the new planet's name, only to learn that the URL was a whimsical tribute to Brown's newborn daughter.)

That same year, Brown, Trujillo, and Rabinowitz collaborated with the engineering team at the Keck Observatory on Mauna Kea, Hawaii, to search 2003 UB313 and three more of the brightest trans-Neptunian objects for satellites. Using the Keck's new Laser Guide Star Adaptive Optics system, which enabled the researchers to view details as precise as those seen from the Hubble Space Telescope, they found S/2005 (2003 UB313) 1—the faint satellite that would come to be called Dysnomia—on September 10, 2005. The team dubbed the satellite Gabrielle, after Xena's television-show sidekick.

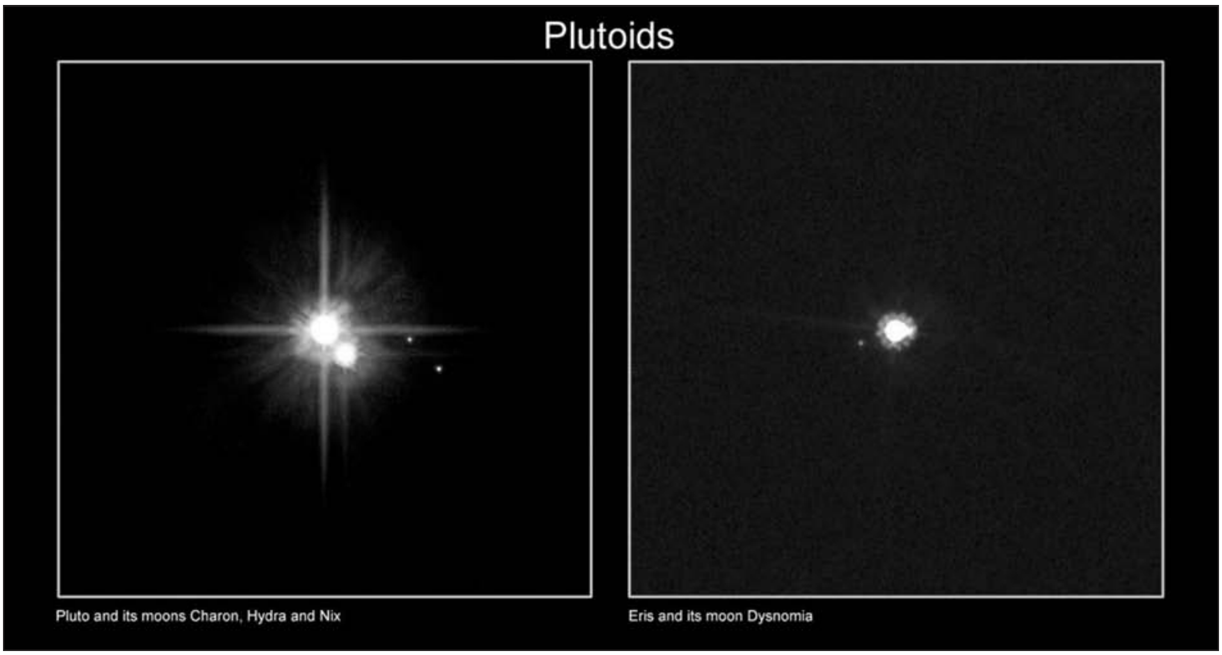
In the summer of 2006, at a meeting of the General Assembly of the International Astronomical Union (IAU), criteria were developed regarding what constitutes a planet, and a new category of "dwarf planet" was established. Under the IAU's new definitions, 2003 UB313 was not a planet but rather a dwarf planet. On September 6, 2006, the discovery team proposed to the IAU that dwarf planet 2003 UB313 be named Eris and its moon be named Dysnomia. The IAU accepted and announced the names one week later on September 13, 2006.

KNOWLEDGE GAINED

The 2005 discovery of 2003 UB313 added fuel to a long-standing and heated debate in the astronomical community over Pluto's status as a planet. The newly discovered object had a greater diameter and was more massive than its distant neighbor Pluto. If Pluto's size and mass were sufficient to qualify it for "planethood," then the new object should likewise be classified as a planet.

The controversy came to a head in August, 2006, at IAU's General Assembly in Prague. At the unusually contentious meeting, members reached the non-unanimous conclusion that a celestial body should be considered a planet if it (1) orbits the Sun (satellites are not planets); (2) is massive enough for self-gravity to shape it into a sphere; and (3) has accreted or scattered other bodies in its neighborhood to clear its orbit. Objects meeting the first two criteria but not the third would be classified as dwarf planets.

Both Eris and Pluto orbit the Sun, and both are spherical, yet neither has cleared the vicinity around its orbit. With the IAU's official acceptance of the new definitions on August 24, 2006, 2003 UB313 ceased to be a possible tenth planet and instead became the largest known dwarf planet. Likewise, Pluto was "demoted" to dwarf planet status after more than seven decades of being regarded as a planet. Ceres was determined to be a dwarf planet that might also be an asteroid.



Pluto and Eris, among the first bodies classified as "plutoids": dwarf planets that orbit the Sun in the icy region beyond Neptune. (International Astronomical Union)

2003 UB313 received its official name, Eris, on September 13, 2006. In Greek mythology, Eris was the goddess of discord and strife. Given the uproar the object's discovery caused among astronomers—and a public accustomed to a nine-planet solar system—it was aptly named. The dwarf planet's satellite was named after the mythological figure Dysnomia, the goddess Eris's daughter and the spirit of lawlessness. The name Dysnomia was also a tip of the hat to Eris's old nickname—the television actress who played warrior princess Xena was Lucy Lawless.

Almost two years later, Eris was assigned to a new subcategory of dwarf planet. On June 11, 2008, the IAU announced that dwarf planets orbiting the Sun at a semimajor axis greater than that of Neptune would be known as "plutoids." Eris and Pluto qualified; Ceres, located in the asteroid belt, did not. In July, 2008, the Kuiper Belt object known as 2005 FY9 received its official name, Makemake, and became the third largest dwarf planet and plutoid.

CONTEXT

Discovery in the early 1990's of small objects beyond Neptune inspired researchers such as Brown to look to the Kuiper Belt and beyond for larger celestial bodies. Brown and Trujillo made their first major trans-Neptunian find in 2002 with Quaoar. In 2003, they discovered a larger object, Sedna, then the remotest solar-system object that had ever been found. Sedna's slower motion across the heavens led the team to look for objects moving at even lower rates, which revealed Eris.

Theoretical models suggest that Eris and other comparatively large trans-Neptunian objects in high-inclination orbits were originally near the inner edge of the Kuiper Belt. When they were subsequently scattered into the outer belt and beyond, they achieved orbits with higher inclinations than objects originating in the outer belt. More massive objects that had their origins in the inner belt may now occupy remote, high-inclination orbits. Researchers are now looking for as-yet-undiscovered large objects orbiting at these high inclinations.

At present, there is no space mission planned to Eris. While the uncrewed New Horizons will conduct a flyby of Pluto, the spacecraft's limited maneuvering capability keeps it from passing close enough to observe Eris.

Karen N. Kähler

FURTHER READING

Brown, M. E., C. A. Trujillo, and D. L. Rabinowitz. "Discovery of a Planetary-Sized Object in the Scattered Kuiper Belt." *The Astrophysical Journal* 635, no. 1 (2005): L97-L100. The journal article in which Eris's discovery team first presented their findings. Explains how the team discovered 2003 UB313 and summarizes the initial physical and orbital characterizations of the object. Includes tables, figures, and references. Technical but straightforward.

_______. "Satellites of the Largest Kuiper Belt Objects." *The Astrophysical Journal* 639, no. 1 (2006): 143-146. The scientific paper in which the discovery team introduced Eris's moon. Describes the survey of 2003 UB313 and three other trans-Neptunian objects for possible satellites. Includes tables, figures, and references. Challenging for the lay reader.

Chang, Kenneth. "Dwarf Planet, Cause of Strife, Gains 'the Perfect Name.'" *The New York Times*, September 15, 2006, p. A20. A brief but informative article on how dwarf planet 2003 UB313 and its satellite came to be named Eris and Dysnomia—and nicknamed Xena and Gabrielle.

_______. "Ten Planets? Why Not Eleven?" *The New York Times*, August 23, 2005, p. F1. Reprinted in *The Best American Science Writing 2006*, edited by Atul Gawande. New York: HarperPerennial, 2006. An engaging profile of team leader Mike Brown, published shortly after the announcement that 2003 UB313 had been found. Includes a chart summarizing five of Brown's major finds, among them 2003 UB313.

Faure, Gunter, and Teresa M. Mensing. *Introduction to Planetary Science: The Geological Perspective*. New York: Springer, 2007. Designed for college students majoring in Earth sciences, this textbook provides an application of general principles and subject material to bodies throughout the solar system.

Excellent for learning comparative planetology.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory text covering the field of astronomy. Contains descriptions of astrophysical questions and their relationships.

Fussman, Cal. "The Man Who Finds Planets." *Discover*, May, 2006, 38-45. A highly accessible account of the discovery of 2003 UB313 (then still popularly known as Xena) and other trans-Neptunian objects, as told by team leader Mike Brown. Includes photographs.

See also: Callisto; Ceres; Dwarf Planets; Enceladus; Europa; Ganymede; Iapetus; Io; Jupiter's Satellites; Lunar Craters; Lunar History; Lunar Interior; Lunar Maria; Lunar Regolith Samples; Lunar Rocks; Lunar Surface Experiments; Mars's Satellites; Miranda; Neptune's Satellites; Planetary Satellites; Pluto and Charon; Saturn's Satellites; Titan; Triton; Uranus's Satellites.

Europa

Categories: The Jovian System; Natural Planetary Satellites

Europa is one of the four "Galilean satellites" that orbit the giant planet Jupiter. Only slightly smaller than the Earth's moon, Europa is covered by a relatively smooth layer of highly reflective fractured ice. Tidal forces exerted by Jupiter cause internal heating on Europa that apparently results in the periodic resurfacing of watery flows, which have over time obliterated most impact craters and other blemishes. Heat flow may be sufficient to maintain a liquid water subsurface layer that could harbor simple life-forms.

OVERVIEW

Europa is one of the four large satellites of the planet Jupiter known as the Galilean satellites after their discoverer, Galileo Galilei. These (according to their distance from Jupiter) are Io, Europa, Ganymede, and Callisto. Jupiter has at least sixty-three satellites, but only the Galilean satellites are large enough to be observed from Earth by small telescopes. With a diameter of 3,138 kilometers, Europa, the smallest Galilean satellite, is slightly smaller than Earth's Moon (3,476 kilometers). By contrast, the largest Galilean satellite, Ganymede, measures 5,260 kilometers in diameter, larger than the planet Mercury (at 4,878 kilometers). Thus, if the Galilean satellites orbited the Sun instead of Jupiter, they would be considered full-fledged planets. Despite its relatively small size compared to its Galilean companions, Europa is nevertheless the sixth largest planetary satellite in the solar system. It is located about 780 million kilometers from the Sun, about 5.2 times the Earth-Sun distance.

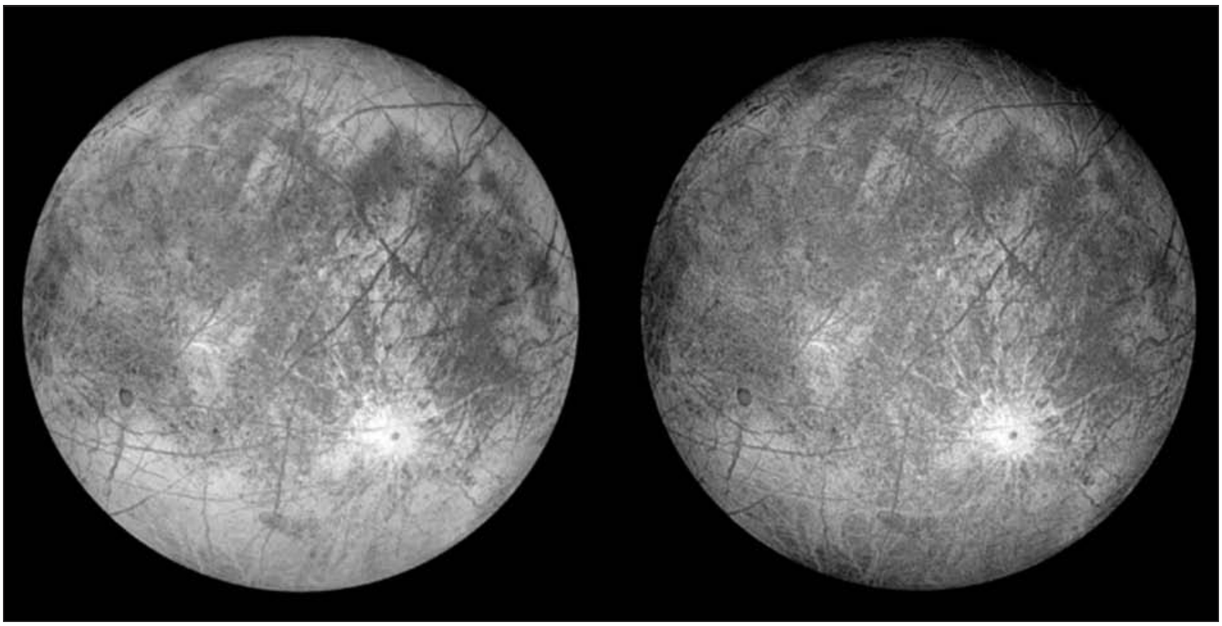
Europa orbits Jupiter at an average distance of 670,900 kilometers; its orbital period (time to complete one orbit) is 3.55 Earth days. Its rotational period around its axis is also 3.55 days, which means that Europa always shows the same face toward Jupiter. The other Galilean satellites and Earth's own Moon follow this 1:1 ratio of orbital to rotational period, termed a "synchronous" relationship.

Galileo discovered Europa and two of the other three large Jovian satellites (Io and Callisto) on January 7, 1610, using a crude homemade telescope. At first he believed the tiny points of light in line with Jupiter were small stars, but later he realized that they in fact orbited Jupiter as if in a miniature solar system. Galileo originally called the moons the Medician planets (after the powerful Italian Medici family) and numbered each satellite with a Roman numeral beginning with the one closest to Jupiter. Europa in this scheme was designated II. Another observer, Simon Marius (Simon Mayr), who claimed to have discovered the Jovian satellites prior to Galileo in November, 1609, but was tardy in publishing his results, later named the bodies as we know them today.

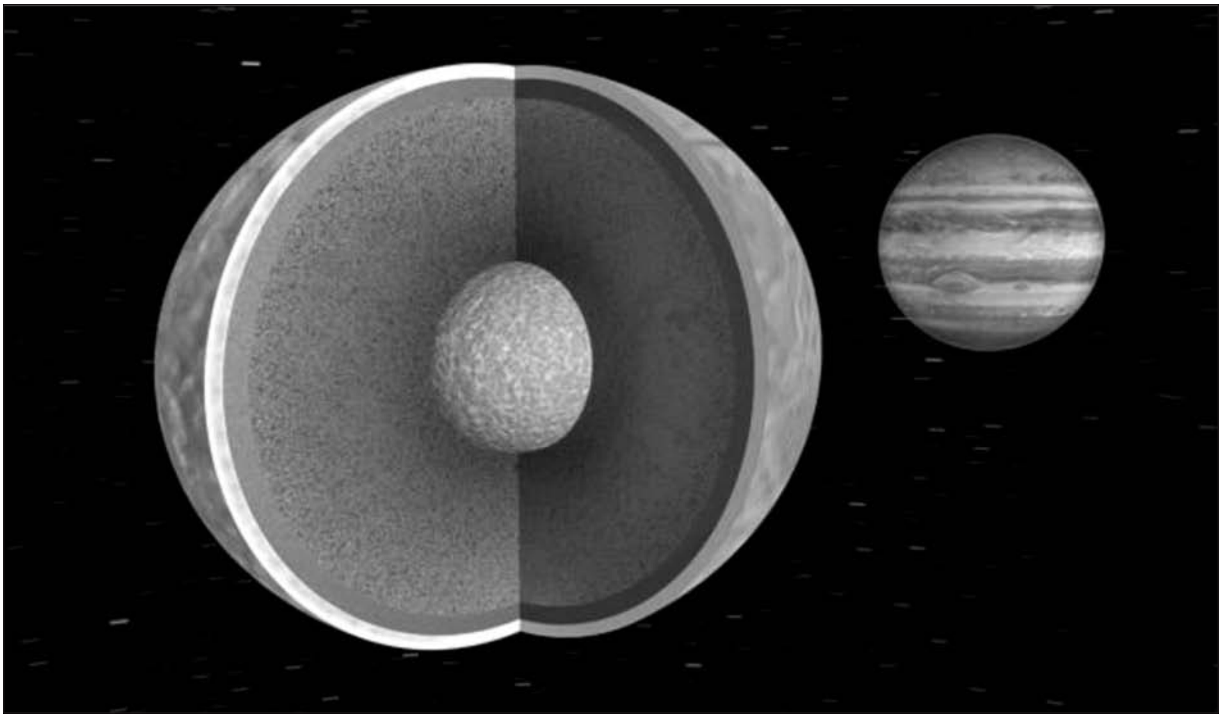
The name Europa comes from a Phoenician princess, one of many mortal consorts of the supreme Greek god Zeus, whose Roman name graces the planet Jupiter. (The other Galilean satellites are similarly named for mythological characters associated with Zeus.) The most intriguing aspect of Europa is its unusual and unique surface. Images beamed to Earth in 1979 by the Voyagers 1 and 2 spacecraft as they flew through the Jupiter system showed a relatively smooth ice ball that some scientists compared in appearance to a fractured, antique ivory billiard ball. The satellite is covered by a globally encompassing shell of water ice, frozen at 128 kelvins, that gives Europa an extremely high albedo. While 64 percent of the light striking the surface is reflected back in all directions (giving Europa an albedo of 0.64), rocky surfaces like that of Earth's moon or Mercury reflect only about 10 percent.

Europa's density, 3.04 grams per cubic centimeter, suggests that most of the planet is composed of rocky silicate material like Earth. The icy surface layer, therefore, must be relatively thin; most estimates lie in the range of 75 to 100 kilometers thick. The surface shows relatively little topographic relief, nothing higher than 1 kilometer, and displays only a few small scattered impact craters, in dramatic contrast to its highly cratered neighbors, the two outer satellites Ganymede and Callisto. Large craters on the order of 50 to 100 kilometers in diameter are virtually absent on Europa but plentiful on Ganymede and Callisto. Most craters on Europa do not exceed about 20 kilometers in diameter. This suggests that Europa's icy surface is relatively young, indicating that resurfacing by liquid ice flows or other processes has covered over any large craters formed during early, heavy meteoroid impacting in the Jovian system. Estimates of the surface age of Europa range from a high of 3.0 to 3.5 billion years old to more recent estimates of only 100 million years. The younger age, if true, suggests significant resurfacing of the planet in the later stages of its history.

In December, 1995, the Galileo spacecraft entered orbit about Jupiter. Over the course of thirty-eight orbits, Galileo not only investigated the giant planet Jupiter but also flew by its many satellites, paying particular attention to Europa, Ganymede, and Callisto. For example, in 1997 Galileo produced images of a large, multiringed impact crater on Europa probably buried beneath the ice crust. Evidence for the crater consists of diffuse, dark, concentric,



The Galileo spacecraft returned these images (both of the same hemisphere, the one on the right enhanced to emphasize details) of Jupiter's satellite Europa, whose surface is primarily water ice. Darker areas are rocky material, lines are crustal fractures, and the bright spot on the lower right is the crater Pwyll. (NASA)



An artist's cutaway showing the probable interior structure of Europa, with an iron core, a rocky mantle, and an outer ocean of salty water capped with ice. (NASA/JPL)

arclike bands and associated fractures that define a structure more than 5,000 kilometers in diameter. The presence of this buried crater shows that the rocky surface below the ice layer was subjected to significant impacting early in Europa's history. It further suggests that the ice crust formed at some later time, probably after heavy meteoroid bombardment had greatly diminished.

The most striking aspects of Europa's surface are the mottled, colored terrains and linear fractures that crisscross most of its globe. Mottled terrains, based on color and subtle topographic expressions, are of two varieties: brown and gray. Brown terrains contain numerous pits and depressions from 1 to 10 kilometers in diameter. Several large "plateaus" occur that range from a few kilometers up to a few tens of kilometers wide and up to nearly 100 kilometers long. Some circular depressions, missing raised crater rims, may represent degraded impact craters. Gray terrains are similar to the brown but are generally smoother and less hummocky. The relationship between the two terrains is unknown, but their differences may result from contrasting ages, degree of surface development, or both. The ultimate origin of these mottled terrains remains unknown, but a reasonable hypothesis is that they represent the effects of hydrothermal upwelling, causing heating and expansion of affected crustal areas. The "nonmottled" areas on Europa are very light in color and have very smooth topographies. These icy plains contain most of the observed linear surface fractures.

Linear features on Europa's surface may extend for thousands of kilometers. They are classified into three categories: (1) dark triple bands, some containing dark outer bands with a white strip down the center, thought to represent icy geyser deposits erupted along the axis of the fractures; (2) older and brighter lineaments that are crosscut by the triple bands and resemble them in some cases; and (3) very young cracks that crosscut the other two fracture types. Detailed analysis of the orientation of these three fracture types indicates that each type shows a distinct orientation that can be correlated with the relative age of the fractures. The data show that the direction of tidal stresses in Europa's crust has rotated in a clockwise direction over time. This observation has been used to suggest that Europa's rotation is not perfectly synchronous. Over time Europa may rotate faster than the synchronous rate, causing the surface to be progressively reoriented relative to tidal forces.

High-resolution images from the Galileo orbiter show places on Europa resembling ice flows in the Earth's polar regions. Large, angular pieces of ice have shifted away from one another, some rotating in the process, but reconstructions show that they fit together like puzzle pieces. This evidence for motion involving fluid flow, along with the possibility of geyser eruptions, shows that the ice crust has been, or is still, lubricated from below by warm ice, or even liquid water. The source of heating to produce this watery fluid is tidal forces by gravitational interaction with massive Jupiter, along with some escaping heat produced by radioactive minerals in the underlying silicate crust.

The Galileo spacecraft's mission was expanded to include the Europa Extended Mission, as it flew a number of close flybys to focus its instrumentation and cameras specifically on Europa. On one close encounter with Europa, Galileo came within 200 kilometers of the icy surface of the satellite. Ultimately, the Galileo spacecraft was purposely directed to plunge destructively into Jupiter's atmosphere on September 21, 2003. The reason for this was to safeguard any possible life-forms on Europa against the plutonium inside the Galileo spacecraft's radioisotope generators in the event that it might have crashed into the satellite.

Europa remains a high-priority location within the solar system for astrobiology studies. With the demise of the Galileo spacecraft, plans were proposed to send another spacecraft, this time to orbit Europa for prolonged and repeated studies. A more ambitious plan arose, called Jupiter Icy Moons Orbiter or JIMO. JIMO would have been the flagship mission of a more extensive program to develop nuclear propulsion as a means to cut down the time of travel between Earth and the rest of the solar system. That program was called Project Prometheus, but after initial funding was granted the National Aeronautics and Space Administration (NASA) was forced to postpone, if not cancel, this futuristic propulsion system in favor of other expenditures arising from the Vision for Space Exploration program under the George W. Bush administration. Returning to Europa with a robotic spacecraft was therefore put on hold for the early portion of the twenty-first century. Data from the Cassini spacecraft in orbit about Saturn revealed aspects about the icy satellite Enceladus and Titan (with its thick atmosphere and organic compounds) that diverted the attention of many astrobiologists away from Europa.

METHODS OF STUDY

Jupiter and its four largest satellites have been studied using telescopes since Galileo first trained his on the system in 1610. Prior to the advent of interplanetary space probes, telescopic observations resulted in a remarkable treasure trove of data on the Galilean satellites.

For example, in the 1920's the astronomers Willem de Sitter and R. A. Sampson succeeded in obtaining reasonably accurate data on their masses. Calculations involved observing how each satellite disturbed the orbits of the others and by noting the nature of the resonant orbits of the inner three (first described by Pierre-Simon Laplace in the late eighteenth century). These resonant orbits dictate that for every one orbit of Io around Jupiter, Europa revolves two times and Ganymede four. This orbital resonance scheme implies a specific ratio for the masses of the bodies, which assisted de Sitter and Sampson in their calculations.

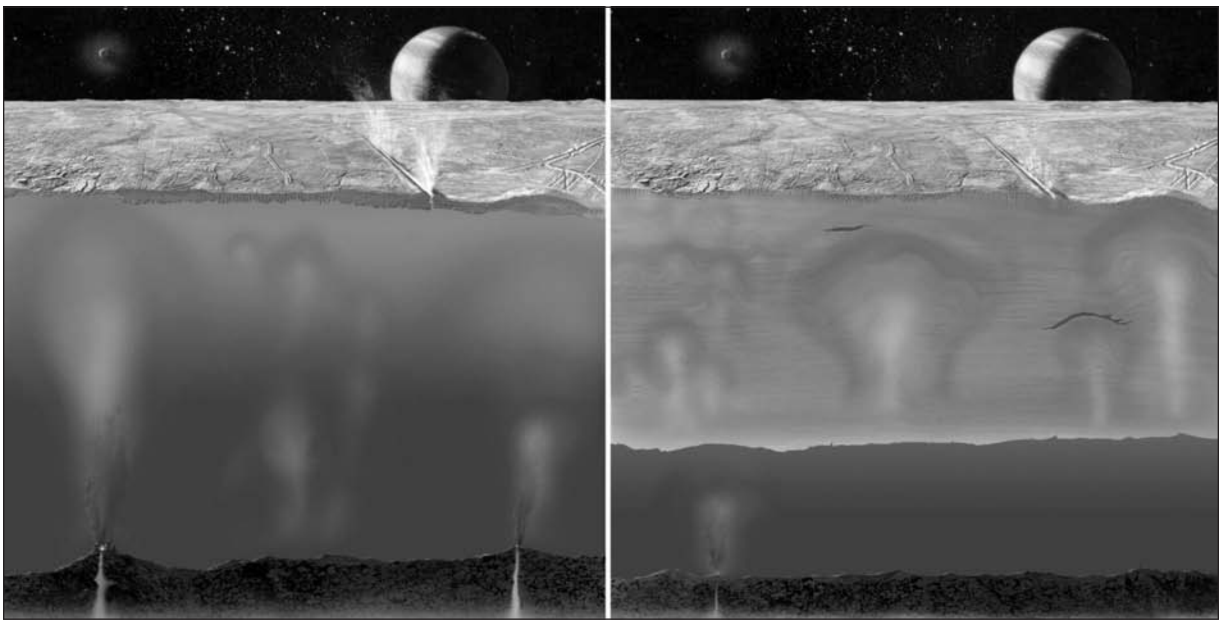
Diameters of the satellites were not accurately known until the advent of stellar occultation studies in the 1970's and later, when spacecraft imaging produced precise values. Prior to that, Europa was described by a popular 1950's-era science text as having a diameter of 1,800 miles (2,880 kilometers), only a bit less than the currently accepted value of 3,138 kilometers.

Although the first Earth-launched space probes encountered the Jovian system in 1973 (Pioneer 10) and 1974 (Pioneer 11), they paid scant attention to the Galilean satellites. The community of planetary scientists at the time viewed the satellites of the gas giant planets in the outer solar system to be nothing but rather boring ice balls. Even on the Voyagers, few planetary science studies were devoted to any of the icy satellites. Only imaging studies of Io and Titan, and to a lesser extent Europa, were planned as major portions of Voyager flyby operations in the Jupiter or Saturn system.

In 1979, however, knowledge of these bodies dramatically expanded as images of all four satellites were beamed back to Earth by Voyagers 1 and 2. The first pictures of Europa showed a previously unknown world, with a highly reflective, smooth surface mottled by brown and tan patches and crisscrossed by a complicated network of curved and straight lines. Four months later, higher-resolution images from Voyager 2 confirmed the presence of even more linear structures, which were interpreted as fractures but having virtually no relief associated with them. In addition to its imaging work, the Voyager probes also made precise measurements of the mass of the Galilean satellites by analyzing the gravitational effects of the planets on spacecraft trajectories, which, combined with improved size determinations, allowed for more accurate calculations of density. Density, in turn, is used to assess planetary composition.

In 1995 the Hubble Space Telescope discovered a thin oxygen atmosphere on Europa; this was later confirmed by the Galileo orbiter during its Europa Extended Mission. The official term for this rarefied atmosphere is a surface-bound exosphere. Hubble used its highly sensitive spectrometers to analyze the energy spectrum of light reflected from the moon's surface. Europa's atmosphere is so tenuous that its surface pressure is only one-hundred-billionth that of Earth. It is estimated that if all the oxygen on Europa were to be compressed to the surface pressure of the Earth's atmosphere, it would fill about a dozen Houston Astrodomes.

The Galileo spacecraft investigated the Jupiter system from late 1995 through 2003. After launching an atmospheric probe into Jupiter itself relatively early in its mission, the Galileo orbiter assumed an elliptical orbit that allowed it to make several close passes to all four Galilean satellites. The resolution of Voyager images of Europa made it possible to view surface features no smaller than about 4 kilometers across. In contrast, Galileo swooped down closer than either Voyager spacecraft, and with its



An artist's cross-section rendering of two likely structures of Europa, with ice overlying a saltwater ocean beneath and heat rising through the rocky mantle, possibly volcanically. (NASA/JPL) more sophisticated cameras it achieved resolutions of around 10 meters per pixel, allowing objects the size of earthly buildings to be discerned. From these high-resolution images scientists have observed evidence of both tensional and compression ridges and have documented features like water-ice geysers, possible ice volcanoes, and jumbled ice flows that resemble puzzle pieces. These observations paint a picture of a dynamic planet in which tectonic faulting and flooding by liquid water occur periodically. The dark color of many surface fractures may result from the injection of water or warm ice mixed with darker silicates that well up into the fractures and freeze. Galileo images have generated renewed interest in the idea that a layer of liquid water exists below the ice or existed some time in the recent past.

Galileo also carried a magnetometer to detect the existence of a planet's magnetic field and to measure its strength. During a December, 1996, pass of Europa this magnetometer detected the first evidence of a magnetic field. Ganymede, the next moon out from Europa, also has a magnetic field. Although it is about four times weaker than that of Ganymede, Europa's field is still of substantial magnitude. Combined with gravity data suggesting a dense core, the Europa magnetic measurements indicate the probable existence of a sizable metallic core, as well as a layered internal structure similar to that of Earth. The magnetic field data also, however, provided constraints on the nature of the water on Europa. The magnetic field could not be explained by assuming pockets of salty water within a crust of ice, but rather required a spherical shell of liquid water.

CONTEXT

The Jovian system has long been of interest to scientists as a possible model analogous to the larger solar system of the Sun and planets. In this model Jupiter is a substitute for the Sun and the Galilean satellites represent the planets, particularly the rocky planets from Mercury to Mars. The considerable masses and stable orbits of the Galilean satellites suggest that they probably originated along with Jupiter during its formation from the gaseous solar nebula. If so, do these satellites show evidence for having evolved in a manner that parallels that of the inner planets of the solar system, including Earth?

In the early 1970's planetary scientist John Lewis pointed out that the densities of the two outer satellites, Ganymede (1.93 grams per cubic centimeter) and Callisto (1.83 grams per cubic centimeter), were consistent with condensation of solar-composition gas (the solar nebula), where water ice is a stable compound. He predicted that these two bodies should be composed of about equal parts water ice and silicate rock, a view generally accepted today. The two inner bodies, Io and Europa, however, have higher densities (3.55 and 3.04 grams per cubic centimeters, respectively) and thus would be expected to contain less in the way of low-density materials like ice, with a density of 1.0 gram per cubic centimeter. In fact, these bodies show evidence of being largely composed of rocky material, with no ice on Io and only a thin crust of ice over an ocean of liquid water on Europa. What processes could have produced such a density distribution?

In the early 1950's astronomer Gerard P. Kuiper suggested that Jupiter had been very hot during its early history. Building on Kuiper's early work, current hypotheses confirm that Jupiter was probably hot enough in its infancy to have forced low-mass, volatile gaseous materials to the outer fringes of the Jovian region, leaving heavier compounds to accrete as planetoids closer to Jupiter. The lighter volatile gas would contain a high proportion of elements that would eventually freeze as ice, compared to denser silicate minerals. These materials would eventually accrete to produce Ganymede and Callisto, while the volatile-poor inner gas would eventually accrete as Io and Europa. Europa, being farther from Jupiter than Io, has more volatiles such as ice than Io, which has virtually no ice on its surface. Io probably also lost much of its volatile component as a result of long-term volcanic activity, the result of tidal heating produced by Jupiter's gravitational field.

In a similar fashion the solar system as a whole exhibits a composition distribution with high-density "rocky" (or terrestrial) planets near the Sun and more volatile-rich bodies in the outer regions. The inner planets show a similar density distribution. Thus, the Jovian satellite system shows that any evolving planetary system on a scale large enough to have a hot central "star" predictably develops a density distribution where low-density, high-volatile planets dominate the outer regions and high-density "rocky" bodies dominate the inner regions.

The study of Europa is important in terms of its possible role as a site of extraterrestrial lifeforms. Images from the Galileo space probe have confirmed ideas spawned after the Voyager flybys that Europa may have a globally encircling layer of liquid water beneath its surface ice layer. Where water exists in liquid form on a planet, life as we know it can theoretically evolve. Europa has joined the ranks of Mars and Saturn's moon Titan as possible sites where primitive life could exist.

Comparative planetology could indeed be actively studied inside the Jovian system by robotic spacecraft. Voyager provided only tantalizing images and raised many questions. Galileo data strongly suggested that the four large Jovian satellites share essentially similar cores in terms of size. Io is volcanically active, essentially taking material deep inside and turning the satellite inside out by resurfacing itself with sulfur and sulfur compounds; therefore Io is devoid of an icy shell. Europa has internal heat that is insufficient to melt the ice cover in total but supplies warmth to a subsurface layer of liquid water. Ganymede and Callisto have thick, icy shells and retain more of their primordial character because they lack significant heating from their cores. Ganymede's magnetic field is something of a paradox in that its character is akin to the dipole field produced by convection within a molten iron core.

Europa once held the primary attention of astrobiologists as the favored place within the solar system for finding some type of life beyond Earth. As a result of extensive robotic investigation of other portions of the outer solar system, it must now share that hopeful spotlight with Titan and Enceladus.

John L. Berkley

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. Cambridge, England: Cambridge University Press, 2007. A comprehensive work about the biggest planet in the solar system, offering a series of articles by recognized experts in their fields of study. Excellent photographs, diagrams, and figures about the Jupiter system and the various spacecraft missions that unveiled its secrets.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. A chapter on the Galilean satellites by Terrence Johnson gives a comprehensive overview of these satellites based on Voyager and Galileo data. The volume is amply illustrated with color images, diagrams, and informative tables. Aimed at a popular audience, this book can also be useful to specialists. Contains an appendix with planetary data tables, a bibliography for each chapter, planetary maps (including Europa), and an index.

Cole, Michael D. *Galileo Spacecraft: Mission to Jupiter*. New York: Enslow, 1999. Provides a full description of the Galileo spacecraft, its mission objectives, and science returns through the primary mission. Particularly good at describing mission objectives and goals. Suitable for a younger audience.

Fischer, Daniel. *Mission Jupiter: The Spectacular Journey of the Galileo Spacecraft*. New York: Copernicus Books, 2001. Suitable for a wide range of audiences, this volume thoroughly explains all aspects of the science and engineering of the Galileo spacecraft. Particularly good discussions about the nature of the Galilean satellites.

Geissler, Paul E. "Volcanic Activity on Io During the Galileo Era." *Annual Review of Earth and Planetary Sciences* 31 (May, 2003): 175-211. The definitive work describing the physics and planetary geology of volcanoes on Io. Provides a complete picture of Voyager and Galileo spacecraft results.

Greely, R. *Planetary Landscapes*. 2d ed. Boston: Allen and Unwin, 1994. This book concentrates on the nature and origin of planetary surface features. It is packed with excellent diagrams, tables, maps, and monochrome images of planets taken by robotic spacecraft.

A chapter on the Jovian system includes a detailed section on Europa, but it is dated prior to the arrival of the Galileo spacecraft into orbit about Jupiter. Contains an extensive reference section and index.

Harland, David H. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. New York: Springer Praxis, 2000. Includes virtually all of NASA's press releases and science updates during the first five years of the Galileo mission, with an enormous number of diagrams, tables, lists, and photographs. Also provides a preview of the Cassini mission. Although the book was published before completion of the Galileo mission unfortunately, what is missing can easily be found on numerous NASA Web sites.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science. The chapter on Jupiter covers all aspects of the Jovian system and spacecraft exploration of it.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs.

Leutwyler, Kristin, and John R. Casani. *The Moons of Jupiter*. New York: W. W. Norton, 2003. Written by the original Galileo program manager, this heavily illustrated work provides discussions of the Galilean satellites and a number of the lesser known Jovian moons. The authors attempt an artful text to accompany the scientific findings, which may or may not be to the taste of all readers.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar-system astronomy suitable for an introductory college course. Filled with supplemental learning aids and solved student exercises. A Web site is available for educator support. Accessible to nonspecialists as well.

See also: Callisto; Ganymede; Io; Jovian Planets; Jupiter's Satellites; Neptune's Satellites; Planetary Satellites; Saturn's Satellites; Uranus's Satellites.

Extrasolar Planets

Category: Planets and Planetology

Indirect methods of observation have revealed the existence of an ever-increasing number of planets orbiting other stars. These extrasolar planets (or exoplanets) have surprised astronomers and led to new theories about planet formation, because they differ from the planets in our solar system. This new evidence suggests the uniqueness of our own solar system.

OVERVIEW

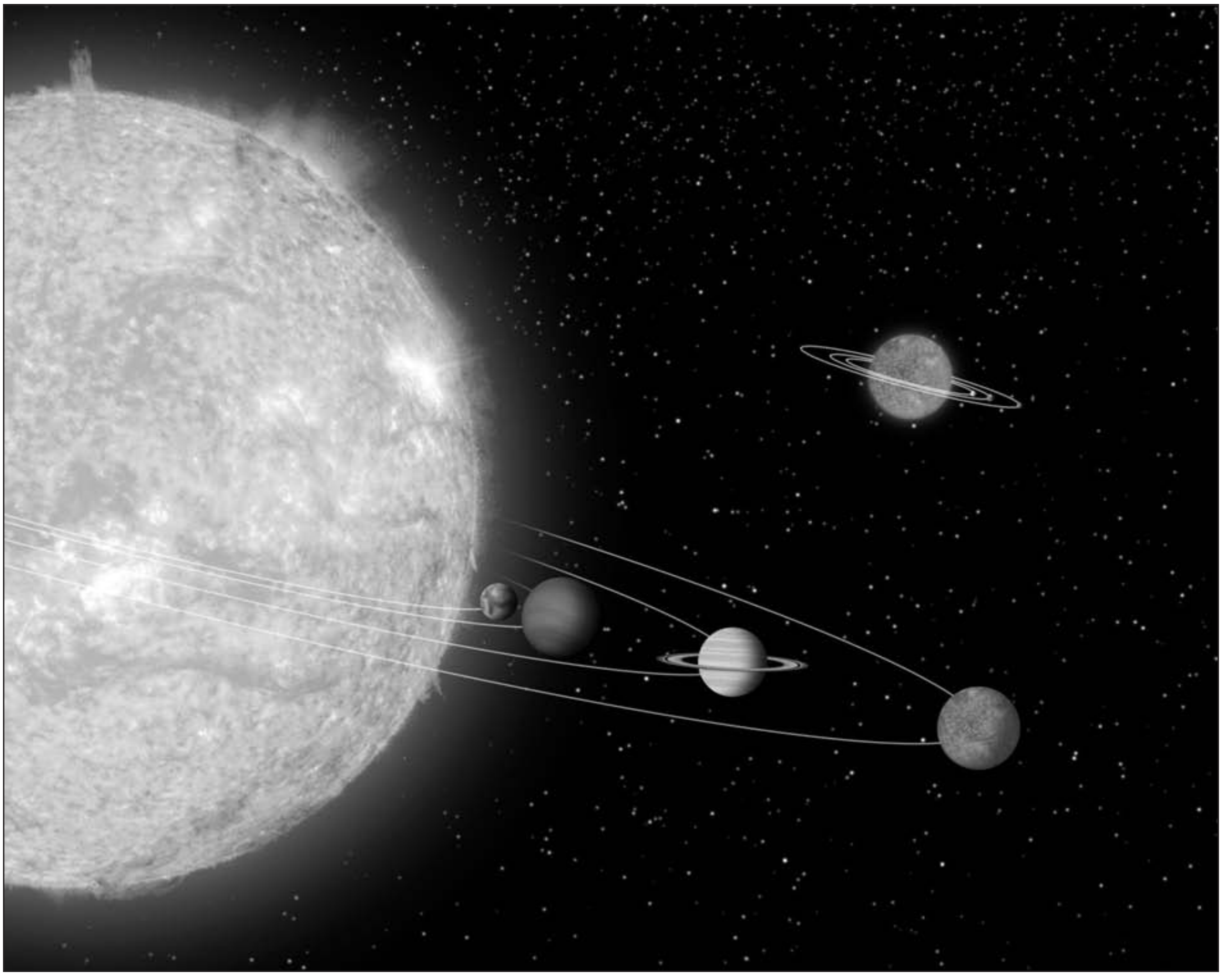
The discovery of extrasolar planets orbiting Sun-like stars has excited the imaginations of astronomers and laypersons alike. If it can be demonstrated that planetary systems are a common occurrence among the billions of stars within our galaxy, the possibility of extraterrestrial life in the universe takes on greater credibility. The idea that intelligent civilizations may exist on other planets could become more compelling.

Early in the twentieth century, spectroscopic evidence from Barnard's star, a nearby red dwarf one-seventh the mass of the Sun, indicated a slight wobble that seemed to imply gravitational interaction by one or two Jupiter-mass planets with decade-long orbits. However, by 1980 further work showed that the wobble of Barnard's star was more likely the result of a companion star too small to observe. The mass of an unseen companion can be estimated from the amount of wobble detected from a visible star.

Double-star systems like Barnard's tend to rotate around their common center of mass in larger orbits than the tiny wobble of a star with a planetary system. Masses between about ten and eighty Jupiter masses usually qualify as brown dwarfs, defined as objects that formed like other stars by gravitational collapse of a dust cloud rather than by accretion from a stellar disk. However, they are too small to sustain the nuclear fusion processes that energize the cores of most stars.

The first confirmed extrasolar planetary system was discovered in 1991, but it was a far cry from a Sun-like solar system that could support life as we know it. Pennsylvania State University radio astronomer Alex Wolszczan was observing a millisecond pulsar (PSR 1257+12) that he and Dale Frail had just discovered using the 305-meter Arecibo radio telescope in Puerto Rico. This pulsar resulted from the collapse of a massive star about a billion years ago. It is now a neutron star that spins 161 times each second, generating a radio pulse about every 6.2 milliseconds. However, Wolszczan found that these pulses varied periodically from the usual high degree of regularity exhibited by other pulsars.

Analysis revealed two periods in the pulse variations from PSR 1257+12: one lasting sixty-six days and the other ninety-five. Wolszczan and Frail proposed that two Earth-size planets orbit the pulsar, gravitationally tugging on it and causing its radio pulses to arrive slightly earlier and then later than expected. Calculations showed that one planet had at least 3.4 Earth masses at an orbital distance of 0.36 astronomical unit (1 AU is the distance from



The Spitzer and Hubble Space Telescopes as well as two Earth-based telescopes found the beginnings of a miniature solar system centered on the star 55 Cancri, located in the Chamaeleon constellation, 500 light-years from our own solar system and depicted in this artist's rendering. (NASA/JPL-Caltech)

Earth to the Sun). The other was at least 2.8 Earth masses and 0.47 AU from the pulsar. By 1994, additional observations revealed a third planet with a period of twenty-three days that had about 0.015 Earth mass located at 0.19 AU.

Planet discoveries about Sun-like stars began in 1995, revealing two new and unexpected types of planetary objects: small-orbit, hot-Jupiter-type planets, and eccentric-orbit Jupiter-like planets. In October of 1995, Swiss astronomers Michel Mayor and Didier Queloz of the Geneva Observatory announced evidence of a companion object orbiting 51 Pegasi (in the constellation Pegasus), which is about 40 light-years away. A new generation of optical instruments and computers revealed a periodic Doppler shift of the light coming from the star. This suggested a tiny wobble caused by a planet of at least 0.46 Jupiter's mass and a period of only 4.2 days in a circular orbit at an orbital distance of 0.05 AU. At this small distance, the planet orbiting 51 Pegasi has a surface temperature of about 1,000 kelvins. While officially named 51 Pegasi b, the planet is informally referred to as Bellerophon.

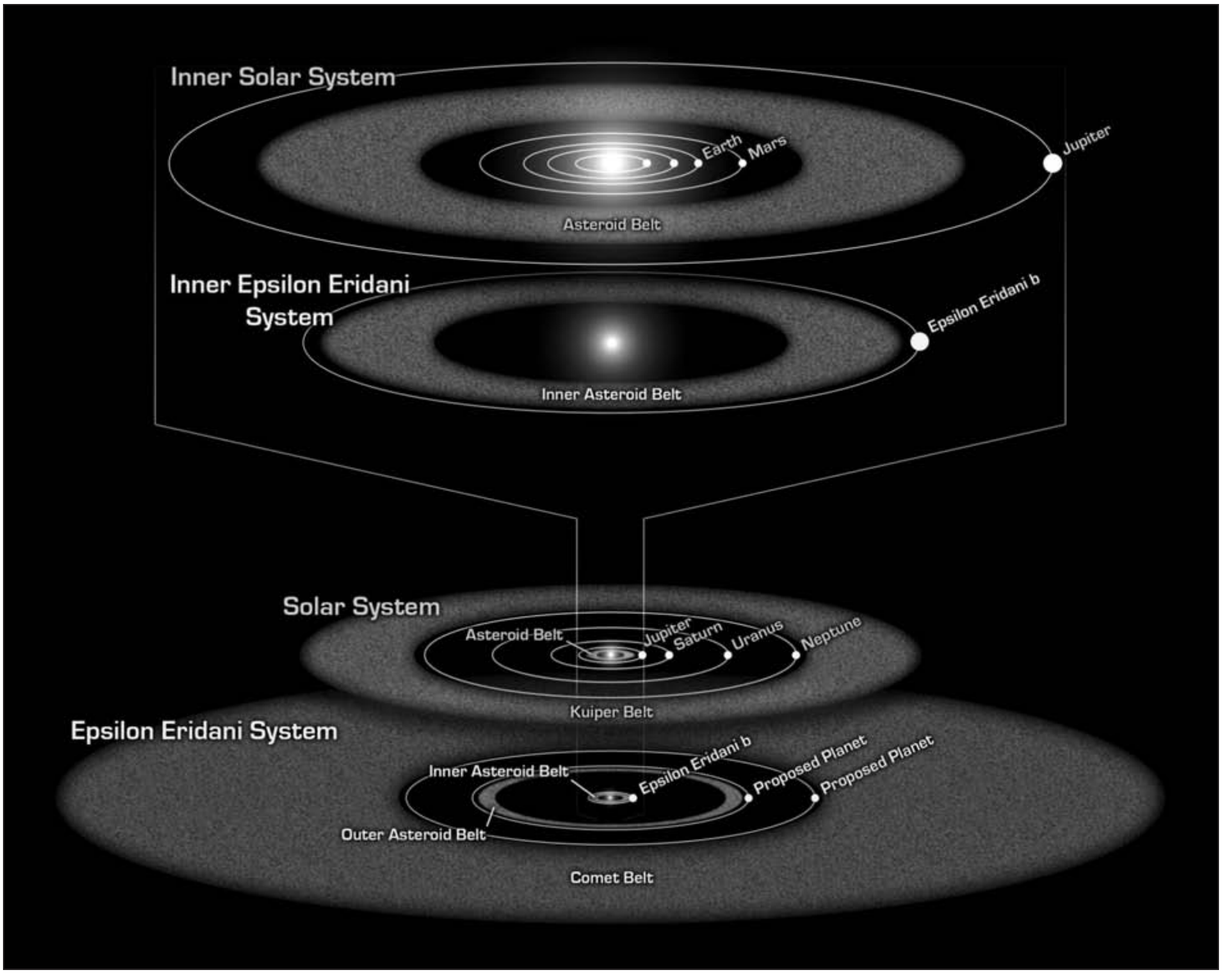
In a 1996 survey of 120 nearby Sun-like stars, Geoffrey Marcy of San Francisco State University and Paul Butler of the University of California, Berkeley, used a refined form of Mayor and Queloz's method to discover six new Jupiter-size planets. The existence of the first two planets, announced in January of that year, were discovered from the tiny wobbles of stars in Virgo and Ursa Major, located 46 and 80 light-years away, respectively. The planet around the star 47 Ursae Majoris has a minimum mass of 2.3 Jupiter masses with an orbital period of 3.0 years and an orbital radius of 2.1 AU (less than half of Jupiter's distance of 5.2 AU). The planet orbiting 70 Virginis has a minimum mass of 6.6 Jupiter masses and a highly eccentric orbit (0.40 eccentricity) of 117 days at an average orbital radius of 0.43 AU. In 2002, Marcy and Butler, along with Debra Fischer, announced their finding of 47 Ursae Majoris c. The planet has an orbital period of 2,594 days and has roughly the mass of Jupiter.

The four other planets included three hot-Jupiter planets similar to 51 Pegasi with nearly circular orbits. At 46 light-years, 55 Cancri has a planet with mass at least 0.8 that of Jupiter and an orbital period of about 15 days, with an orbital radius of 0.11 AU. At 60 light-years, Tau Bootis has a planet with a minimum mass of 3.87 Jupiter masses, a period of 3.3 days, and an orbital radius of only 0.046 AU. Located 55 light-years away from Earth, Upsilon Andromedae has a planet with mass at least 0.68 that of Jupiter, an orbital period of 4.61 days, and an orbital radius of 0.06 AU. Eleven years later, in 2007, astronomers confirmed the existence of two other planets in the Upsilon Andromedae system. The middle planet has an approximate orbit of 242 days and is at least twice the mass of Jupiter. The outermost planet, at 2.5 AU, is at least four times as massive as Jupiter, with an orbital period between 3.5 and 4 years.

Marcy and Butler also announced a possible second planet orbiting 55 Cancri with a minimum mass of about 5 Jupiters, an orbital period of about 20 years, and an orbital radius of 5 to 10 AU. In 2002, 55 Cancri d was found orbiting the star at a distance of about 5 AU, with a mass 4.8 times that of Jupiter. At that time, a third planet—named 55 Cancri c and having a mass roughly that of Saturn's and a highly eccentric orbit—was also speculated to exist. In 2004, 55 Cancri e was discovered. This planet is about the size of Neptune. It is either a small gas giant or a large terrestrial planet. At this time, the existence of 55 Cancri c was confirmed.

The following year astronomer Jack Wisdom questioned the accuracy of naming 55 Cancri e as an exoplanet. After reexamining available data, Wisdom believed that planet e has an orbit of 261 days instead of the proposed 2.8-day orbit. In 2007, 55 Cancri f was announced to be about half the mass of Saturn and to have an orbital period of 260 days, placing it in the habitable zone of the star. 55 Cancri f is not believed to contain life, but hypothetically any satellites the planet has could contain at least microbial life-forms.

Evidence for the nearest planetary system was also announced in 1996 by George Gatewood of the University of Pittsburgh. He collected photometric data on many of the nearest stars with the 30-inch refractor telescope at Allegheny Observatory. The dim red dwarf star Lalande 21185, the sixth nearest star to the Sun



A comparison of the Epsilon Eridani system, the planetary system closest to Earth's, with Earth's solar system. (NASA/JPL-Caltech)

at 8.2 light-years away, appears to have two Jupiter-size planets in orbits similar to those of the gas giants in our solar system. Gatewood analyzed data from fifty years of photographic observations and eight years of photoelectric measurements, revealing tiny accelerations of the star that suggest one planet of about 0.9 Jupiter mass with a period of about 5.8 years in a circular orbit with an orbital radius of about 2.2 AU (similar to the asteroid belt) and a second planet of about 1.1 Jupiter mass with a period of about 30 years in a circular orbit with a radius of about 11 AU (similar to Saturn). A third, unconfirmed planet may orbit beyond these two Jupiter-like planets. The proximity of Lalande 21185 suggests the possibility of eventually capturing an image of its planets with the Hubble Space Telescope.

Two more eccentric planets were announced in 1997. A group of Harvard astronomers led by David Latham discovered an object in 1988 with a mass of at least 9 Jupiters orbiting the star HD 114762 in an 84-day eccentric orbit that varies from 0.22 AU to 0.46 AU (0.35 eccentricity). For eight years, this object was classified as the smallest known brown dwarf, but after Marcy and Butler announced the 70 Virginis planet with a very similar eccentric orbit varying from 0.27 to 0.59 AU and a minimum of 6.5 Jupiter masses, the companion of HD 114762 appeared to qualify as a possible planet.

A third eccentric planet has by far the greatest eccentricity (0.67) of any known planet. Discovered by Marcy and Butler, it was also independently discovered by William Cochran and Artie Hartzes of the University of Texas at Austin. The planet orbits the star 16 Cygni B, a near solar twin that belongs to a triple star system 100 light-years away. It has a mass of at least 1.5 Jupiters and a 2.2-year orbit that varies between 0.6 and 2.8 AU, giving it wild seasonal variations.

Another hot-Jupiter planet orbiting the star Rho Coronae Borealis appears to fill a gap between the very close 51 Pegasi-like planets (less than 0.11 AU) and the 47 Ursae Majoris planet (2.2 AU). It was discovered in 1997 by a Harvard University team of astronomers led by Robert Noyes and has a nearly circular orbital radius of 0.23 AU, a period of 39.6 days, and a minimum mass of 1.1 Jupiter masses. Given the existence of giant planets with orbits from 0.046 AU (Tau Bootis) to 2.2 AU in a relatively continuous distribution, planet formation theories face dramatic challenges, especially since existing theories predict that Jupiter-size planets cannot form within 5 AU of their host stars.

In 2005, Osiris (HD 209458b), an exoplanet of 0.69 Jupiter mass orbiting at 0.046 AU around its parent star, was directly detected using the Spitzer Space Telescope. The telescope was able to observe infrared light coming from the planet itself. Astronomers noted differences in the light being produced when the planet was transiting in front of the star and also when the planet was blocked by the star. By factoring out the star's constant light, scientists were able to isolate the planet. From this, they were able to estimate the temperature of Osiris to be at least 1,023 kelvins.

HD 189733b was discovered in October, 2005, at a distance of about 63 light-years from Earth. The planet is considered a hot Jupiter, with a mass 15 percent greater than Jupiter's and an orbital period of just 2.2 days. In 2007, using the Spitzer Space Telescope, astronomers in Switzerland detected water vapor within the atmosphere of that planet.

The first "super Earths" were found in 1991 orbiting a pulsar. The two planets were only four times the mass of the Earth, too small to be considered gas giants. The general definition of a super Earth is a terrestrial exoplanet that has a mass between one and ten times that of the Earth. Other scientists say a super Earth must be five to ten times the mass of the Earth. In 2007, Stéphane Udry and his team announced the finding of two super Earths orbiting around Gliese 581. Both planets are within the habitable zone (the area where liquid water could potentially exist) of their star. Gliese 581 c orbits the star at 0.073 AU and has about five times the mass of the Earth.

In 2007, the Mount John University Observatory in New Zealand discovered the smallest extrasolar planet to date. The planet, MOA-2007-BLG-192LB, is only 3.3 times as massive as the Earth and orbits a brown dwarf. The first group of super Earths within the same planetary system was found orbiting HD 40307 in 2008. The three planets are all orbiting the star at a distance less than that of Mercury from the Sun, with nearly circular orbits.

METHODS OF STUDY

Detecting extrasolar planets from Earth is extremely difficult, requiring a new generation of computers and optical instruments. Planets are about a billion times fainter than their host star, making them virtually undetectable by direct methods. An indirect method involves searching for a tiny wobble in the motion of a star as it and any companions orbit about their common center of mass. Although the gravitational interaction between a star and a planet-sized object is too small to observe directly, the radial velocity (back and forth along the line of sight) alternately increases and decreases the wavelength of light from the star, causing an alternating Doppler shift toward the red and then blue end of its spectrum.

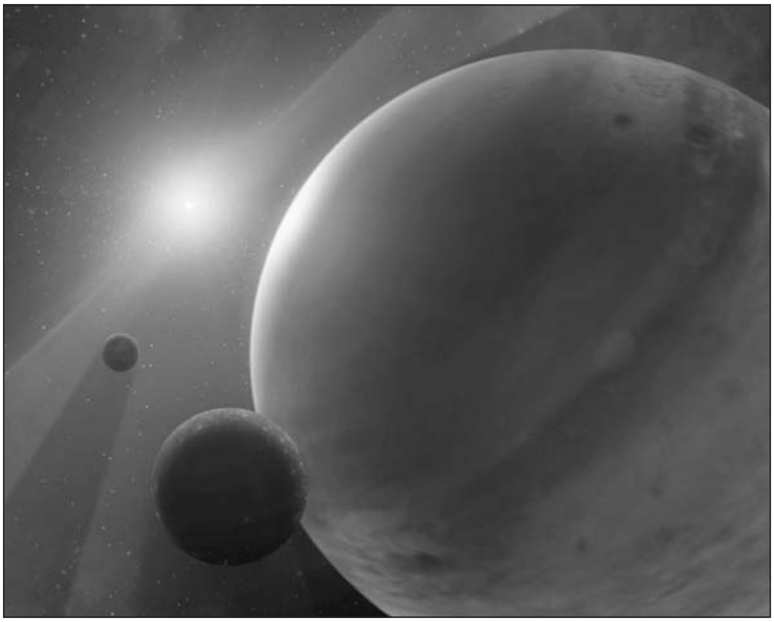
The velocity of a star can be determined from the magnitude of its Doppler shift. The shift in wavelength due to a Jupiter-size planet is only one part in ten million. An absorption cell (consisting of a bottle of iodine vapor placed near the focus of the telescope) absorbs certain known wavelengths of light, producing dark lines in the spectrum that act as a reference for measuring the Doppler shift accurate to within one part in a hundred million. These shifts are recorded by sending light from a star into complex spectrometers consisting of prisms, mirrors, and gratings costing several million dollars.

Periodic variations in the Doppler shift reveal the period of a planet's orbital motion. The velocity of the star and the period of its motion can be analyzed to determine the radius of the orbit (from Kepler's law) and the minimum mass of the planet (from Newton's laws). However, the unknown inclination of its orbit allows for a larger wobble than its apparent radial motion and thus a larger possible mass by a factor of about two. The periodic variation in Doppler shift also reveals the shape of the orbit, since a circular orbit produces a perfect sine wave, while an eccentric orbit produces an irregular variation that can be analyzed by computer to determine the orbital shape.

Using these methods, Marcy and Butler detected radial motions accurate to within ±3 meters per second, compared to at least 10 meters per second required to detect a planet. Since Jupiter, which contains most of the mass of the solar system at 318 times the mass of the Earth, causes the Sun to move at a speed of up to 12.5 meters per second, Jupiter-size planets can be readily detected. Most of the new planet discoveries have been based on stars wobbling at speeds between about 10 and 300 meters per second. Planets much smaller than Jupiter cannot be detected accurately with this method, and those with periods of several years require that data be collected over a long enough time span to determine their periodic variations.

Marcy and Butler began collecting Doppler-shift data in 1987 for their survey of 120 Sun-like stars, using Lick Observatory's three-meter telescope; but it was the computer methods used by the Swiss in their discovery of the 51 Pegasi planet that finally yielded results. Their first discoveries resulted from running six computers day and night at the University of California, Berkeley, to analyze data from sixty stars. These methods revealed a variety of planets that shocked astronomers because their orbits were so unexpected. Hot Jupiters and eccentric orbits have initiated a new generation of theories about planetary formation and the uniqueness of our solar system.

The Spitzer Space Telescope (SST) was launched in 2003. It consists of three main instruments: the Infrared Array Camera (IRAC), an infrared camera that operates simultaneously on four different wavelengths; the Infrared Spectrograph (IRS), a spectrometer able to observe at four wavelengths; and the Multiband Imaging Photometer for Spitzer (MIPS), which is made up of three different far-infrared detector arrays. In 2005, the SST was the first telescope to detect light from exoplanets HD 209458b and TrES-1. However, the light was not turned into actual images.



An artist's concept of a young solar system in which gas giant planets are forming. (NASA/JPL-Caltech/T. Pyle, SSC)

In 2004, a group of astrophysists in France captured the first photograph of an extrasolar planet orbiting a brown dwarf. The planet appeared only as a small red dot. It is speculated to have a mass two to five times that of Jupiter, but it orbits the star at a distance greater than Pluto's average distance from the Sun. This "exoplanet" did not form from an accretion event the way scientists currently believe planetary formation occurs. Also, a brown dwarf would not have enough material to form a Jupiter-sized planet, especially at such distances. Because of these objections, many astronomers do not consider the photograph to be of a real planet.

With the number of known extrasolar planets totaling more than three hundred, the next challenge is to capture a photograph of an actual exoplanet. Two of the programs dedicated to imaging a planet are located in Chile. In 2007, the Gemini South Observatory installed the first optics system specially designed to photograph exoplanets. In 2008, it started a two-year program to conduct a survey of young stars using its Near-Infrared Coronagraphic Imager (NICI). The NICI consists of a coronagraph and two cameras that can simultaneously photograph the star and its surroundings at two different infrared wavelengths. The two photographs would then be subtracted, leaving behind an image of the planet. This method will also help eliminate false planets that are actually background stars or stray starlight.

A possible first photograph of an extrasolar planet (1RXS J160929.1-210524 b) was announced in September, 2008. The image was taken using the Gemini North Telescope on Mauna Kea in Hawaii. The extrasolar planetary system is about 500 light-years away from the Earth. 1RXS J160929.1-210524 b has a mass eight times that of Jupiter and orbits its star at a distance of 330 AU. Some scientists are skeptical of its being an extrasolar planet because of its great distance from its star (Neptune, by comparison, is only 30 AU from the Sun).

CONTEXT

The discovery of extrasolar planets may seem at first to offer new hope for the existence of planetary systems like Earth's solar system that could support extraterrestrial life. However, the unexpected nature of these planets has raised new challenges for planet formation theories and new doubts about the possibility that any of them might harbor life. Pulsar planets were probably formed from the remnants of a companion star during a supernova explosion that produced a spinning neutron star, and they are bathed with high-energy radiation that would make life impossible. The other new planets orbit more Sun-like stars but have either extremely small or highly eccentric orbits that also make them unlikely candidates for life. Evidence so far seems to indicate that our solar system is highly unusual, if not completely unique, in harboring water-based life-forms.

Joseph L. Spradley,
updated by Jennifer L. Campbell

FURTHER READING

Casoli, Fabienne, and Thérèse Encrenaz. *The New Worlds: Extrasolar Planets*. New York: Springer Praxis, 2007. The author discusses the history of the search for extrasolar planets, techniques used, and early discoveries. Also examines research and discoveries since 1996, during which the number of extrasolar planets increased to more than three hundred. The authors speculate about the possibility of life on these planets, as well as what they can teach us about planetary formation.

Dick, Steven J. *The Biological Universe: The Twentieth-Century Extraterrestrial Life Debate and the Limits of Science*. Cambridge, England: Cambridge University Press, 1999. Chapter 4, "Planetary Systems: The Limits of Theory," provides a good history of the search for extrasolar planets before 1995, with several illustrations.

Dvorak, Rudolf, ed. *Extrasolar Planets: Formation, Detection, and Dynamics*. Weinheim, Germany: Wiley-VCH, 2008. This work explains not only how extrasolar planets are detected but their formation, dynamics, and atmospheres as well. It also discusses habitable zones, along with plans to locate and study new extrasolar planets.

Goldsmith, Donald. *Worlds Unnumbered: The Search for Extrasolar Planets*. Illustrations by Jon Lomberg. Sausalito, Calif.: University Science Books, 1997. The first book to discuss the new extrasolar planet discoveries in detail, including theories of formation, methods of observation, and possibilities of life. Includes several color plates and an index.

Jet Propulsion Laboratory, National Aeronautics and Space Administration. *Planet Quest:*

New Worlds Atlas. http://planetquest.jpl.nasa.gov/atlas/atlas_index.cfm. A searchable database listing all currently known extrasolar planets and all known data about the planets and their stars, including scale diagrams comparing them to bodies in our solar system. Regularly updated by NASA's Jet Propulsion Laboratory in California.

Mammana, Dennis, and Donald McCarthy. *Other Suns, Other Worlds? The Search for Extrasolar Planetary Systems*. New York: St. Martin's Press, 1996. A comprehensive history of the search for extrasolar planets through 1995, with several plates, including one describing the January, 1996, discoveries of Marcy and Butler.

Ollivier, Marc, et al. *Planetary Systems: Detection, Formation, and Habitability of Extrasolar Planets*. New York: Springer, 2008. Explores the information known about the three hundred exoplanets currently known, including their sizes, atmospheres, locations, and habitability. The authors also explain the current and possible future methods for detecting exoplanets, as well as the importance of studying young star systems. For the enthusiast.

Scharf, Caleb. *Extrasolar Planets and Astrobiology*. Herndon, Va.: University Science Books, 2008. An advanced technical explanation of extrasolar planets designed for undergraduates or first-year graduate students. Includes helpful chapter summaries and problem sets.

See also: Extrasolar Planets: Detection Methods; Extraterrestrial Life in the Solar System; Infrared Astronomy; Jupiter's Interior; Planetary Formation; Protostars; Solar System: Origins.

Extrasolar Planets: Detection Methods

Category: Planets and Planetology

The search for extrasolar planets, or exoplanets, beyond our solar system and orbiting other stars has yielded several hundred such objects. Detection methods have been mostly limited to finding large Jupiter-like planets, but as such methods improve there is special interest in finding small Earth-like exoplanets in habitable zones that could support extraterrestrial life.

OVERVIEW

Interest in planets orbiting other stars has a long history, but highly sensitive detection methods are required and the first extrasolar planets were not confirmed until the 1990's. After Nicolaus Copernicus introduced his heliocentric theory of a Sun-centered planetary system in the sixteenth century, astronomers began to realize that space might be endless, with an infinite number of stars. At the end of the century, Giordano Bruno proposed that the stars were also suns with their own planets and suggested that there might be an infinite number of other populated worlds.

In 1855, an astronomer at the Madras Observatory of the East India Company claimed that a revolving double-star system in the constellation Ophiuchus (Ophiuchi 70) had orbital anomalies in its eighty-eight-year period that were probably caused by a planet around one of the stars. This claim was repeated in 1896 by American astronomer Thomas See, who calculated that the anomalies were caused by a planet with a thirty-six-year period. These claims were refuted in 1899 by Forest Moulton, who proved that such a three-body system would be highly unstable.

In the 1960's Peter van de Kamp of the Swarthmore College Observatory claimed to have discovered possible planets around Barnard's star, also in Ophiuchus and the second closest star to our Sun. This faint star, which is moving rapidly toward the Sun at about

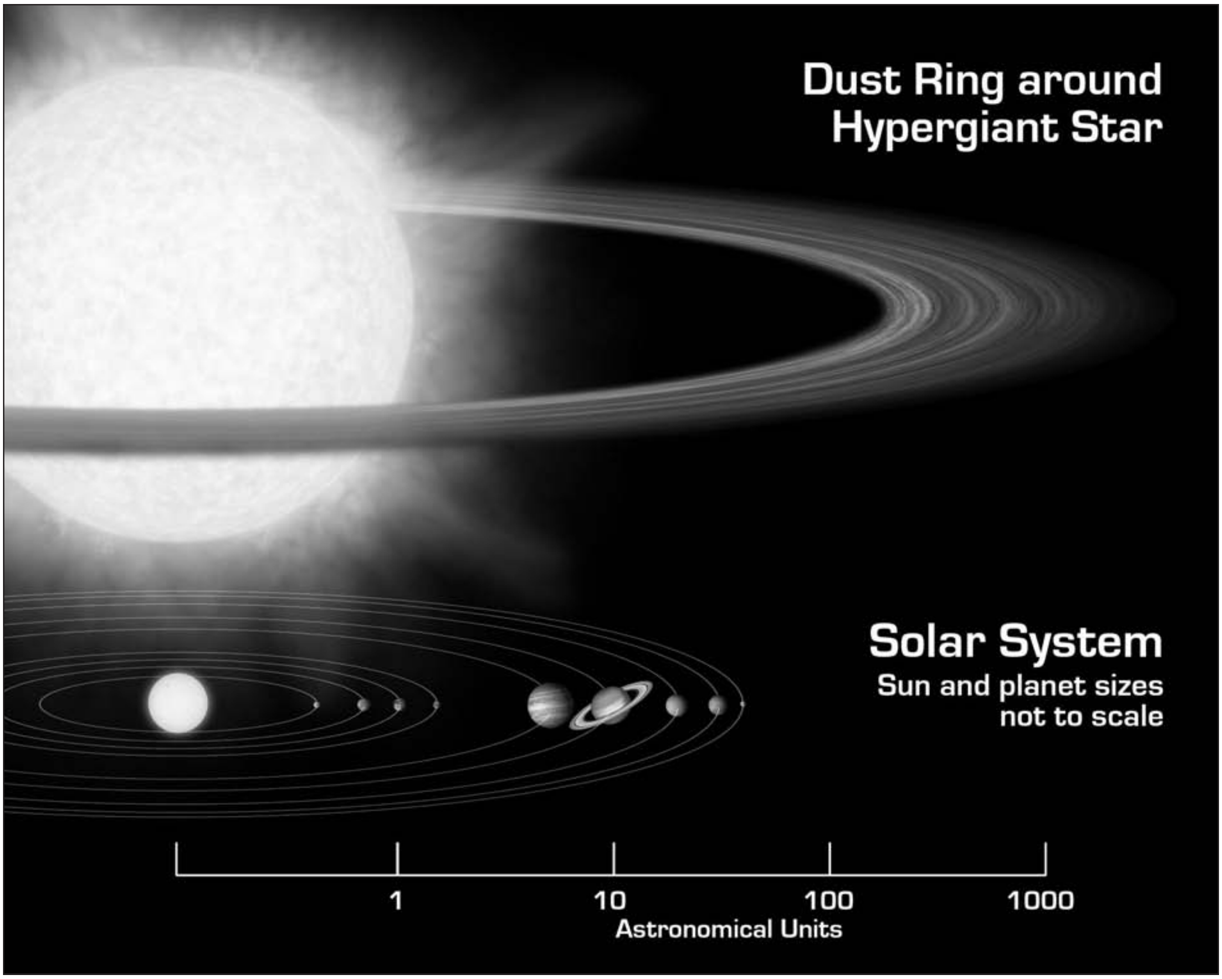
140 kilometers per second, appeared to have a tiny wobble in its motion consistent with two Jupiter-size planets. However, this apparent wobble was not found by other observers, and it was later shown to be caused by lens adjustments. Since the reflected light of a planet is much dimmer than its parent star, most extrasolar planets have been discovered by indirect detection methods beginning in the 1990's. However, since 2004 astronomers using the European Southern Observatory's Very Large Telescope array in Chile have produced direct images of several brown dwarf stars with companions, and in 2005 one of these was confirmed as a planet with a mass several times larger than Jupiter's. Six indirect methods have been used to discover most of the known extrasolar planets, and other detection methods continue to be developed.

Three search methods try to detect the tiny elliptical wobbling of a parent star caused by the gravitational influence of an orbiting planet: astrometry, pulsar timing, and radial-velocity detection. The oldest method of searching for extrasolar planets is by astrometry, which requires precise measurements of tiny variations in the position of a star. Several astrometric discoveries of exoplanets were claimed in the 1950's and 1960's, but none was confirmed. Such movements are probably too small to observe with ground-based telescopes but were demonstrated with the Hubble Space Telescope in 2002. Future plans to search with the National Aeronautics and Space Administration's (NASA's) Space Interferometry Mission may reveal many new planets by astrometry. This method is most sensitive to planets with large orbits and long periods, complementing other methods that are more sensitive to small orbits with short periods.

The first confirmed discovery of an exoplanet used a pulsar timing detection method. Pulsars are rapidly rotating neutron stars that emit rapidly pulsed radio waves at highly regular rates matching the rotation rate. In 1992, radio astronomers Alexander Wolszczan and Dale Frail detected slight periodic changes in these millisecond pulse rates and recognized that they were caused by wobbling of the star due to three planets. This method is so sensitive that it can detect planets smaller than those detectable by any other method—down to a tenth of the Earth's mass—but is limited by the limited number of known pulsars. Although the existence of such small pulsar planets is of interest, they do not offer the possibility of life as we know it since a neutron star emits radiation deadly to such life.

The radial-velocity or Doppler method detects back-and-forth variations in the wobbling of a star and has accounted for the majority of exoplanet discoveries. These radial motions relative to the Earth cause shifting of the star's spectral lines due to the Doppler effect, which decreases and increases the wavelength of the light as the star moves toward and away from the Earth respectively. Modern spectrometers can detect velocity variations down to about 1 meter per second, including the High Accuracy Radial Velocity Planet Searcher (HARPS) spectrometer at the European Space Agency's 3.6-meter telescope in Chile. This method requires high precision and is limited to nearby stars within about 160 light-years. It is most sensitive to large planets with short periods, known as "hot Jupiters" because of their proximity to the Sun, while longer periods require many years of observation. From the period of the planet, the orbital radius can be found. Velocity variations permit an estimate of a planet's minimum mass, which can be considerably larger if the orbit is highly inclined to the line of sight.

A few exoplanets have been detected by three more problematic but developing methods. The transit method measures the tiny dimming of a star when a planet passes in front of it. This method can reveal the size of the planet and can be combined with data from the Doppler method to find its true mass and density. The gravitational microlensing method is based on observations of a star whose gravitational field functions like a lens, focusing light from a distant star directly behind it in the same line of sight. Anomalies in the lensing light curve can reveal planets orbiting Sun-like foreground stars down to the size of the Earth. However, the lensing observations cannot be repeated when the chance alignment of two such stars changes. The circumstellar disk method analyzes the infrared radiation emitted by dust disks that sur-



An illustration of a giant star surrounded by a dust ring, similar to those surrounding the hypergiant stars R 66 and R 126, located in the Large Magellanic Cloud, next-door to the Milky Way, and discovered by the Spitzer Space Telescope. Such dust disks may represent either the seeds of a planetary system or the debris left over after their demise. (NASA/JPL-Caltech/R. Hurt, SSC)

round many stars. Images of dust disks have been obtained by the Hubble and Spitzer space telescopes, and some of these have features that imply the presence of planets.

Three detection methods that may be used in the future are of more specialized interest. The eclipsing binary method looks for anomalies in the light variation as one star of a binary system passes in front of its companion, giving evidence of planets in the system. In the orbital phase method, future space telescopes may be able to detect light variations due to the reflected light from planets that produce phases like that of the Moon. The polarimetry method studies the tiny fraction of polarized light from a star if it passes through the atmosphere of an orbiting planet.

KNOWLEDGE GAINED

Discoveries of extrasolar planets since the first confirmed pulsar planets have increased rapidly and have added much to our knowledge of the varied nature of such planets. The first confirmed planet orbiting a Sun-like star (51 Pegasi) was discovered by the radial-velocity method in 1995 by Michel Mayor and Didier Queloz of the University of Geneva in Switzerland. They were surprised to find that its period was only about 4.2 days and that its orbital radius was much less than that of Mercury. These measurements gave a minimum mass of about half that of Jupiter, or at least 150 times the Earth's mass.

The discovery of a planet around 51 Pegasi was confirmed by a California team led by Geoffrey Marcy, who used the radial-velocity method to discover nearly two-thirds of about three hundred possible extrasolar planets found over the next dozen years. These are mostly hot Jupiters whose masses are assumed to be less than the limit of about thirteen Jupiter masses that distinguishes a planet from a brown dwarf star. Among their achievements was the discovery of the first multiple-planet system, with three Jupiter-size planets around a Sun-like star, and the first transit detection of a planet previously discovered by the radial-velocity method, giving its actual mass and confirming that it was a planet. About twenty multiple-planet systems have been found, and four pulsar planets are known around two separate pulsars.

Among approximately three hundred extrasolar planet candidates, the vast majority have Jupiter-size masses. Since most of these were discovered by the radial-velocity method, only their minimum masses are known and the actual masses of some could eventually show that they are brown dwarfs. About sixty extrasolar planets have been confirmed by various methods that have determined their actual masses. Early discoveries were mostly hot Jupiters with large masses situated very close to their parent star and with very short periods. They seem to defy theories of planet formation based on studies of our solar system, which suggest that large gas planets form farther from their parent star than the smaller rocky planets. However, it is now believed that hot Jupiters probably formed farther out and then migrated in due to larger amounts of dust in their circumstellar disks.

Frequent observations of these hot Jupiters appear to be the result of a selection effect, since such planets are easier to detect over shorter time periods because of the larger and faster wobbling of their parent stars. As detection methods have improved, many such large planets have been found with larger orbits comparable to those of Jupiter and Saturn. Since 2004, a few Neptune-sized planets have been discovered with masses in the range of about seven to fourteen Earth masses. Most exoplanets have much more eccentric orbits than those in Earth's solar system, which is not due to an observational selection effect and is still a major puzzle for astronomers.

CONTEXT

The primary interest in extrasolar planets arises from the possibility of finding Earth-like planets in the habitable zone about their parent stars, where water and thus life might exist. The hot Jupiters and eccentric orbits of most of the exoplanets found so far appear to have little possibility of life. However, a few discoveries are beginning to reveal Earth-like qualities and similarities to our solar system. About seven exoplanets have now been discovered by gravitational microlensing, including a two-planet system found in February of 2008 similar to the Jupiter-Saturn system, which plays the important role of sweeping up errant comets and asteroids that might otherwise make life on Earth impossible.

In 2008 the Optical Gravitational Lensing Experiment (OGLE) at Princeton University found two rocky "super-Earth" exoplanets with masses of only five and three times the Earth's mass, the latter orbiting a brown dwarf, suggesting that it might be a planet covered with water. In May of 2008 several dozen possible new super Earths were announced by an MIT group in Chile using the radial-velocity method with the HARPS spectrometer.

Several planned space missions offer the possibility of much better exoplanet detection from outside the Earth's atmosphere. NASA's Kepler mission is expected to use the transit method with a special space photometer that can simultaneously scan a hundred thousand stars and should be able to detect Earth-sized planets and their statistical frequency around Sun-like stars.

Joseph L. Spradley

FURTHER READING

Casoli, Fabienne, and Thérèse Encrenaz. *The New Worlds: Extrasolar Planets*. New York:

Springer Praxis, 2007. Provides a comprehensive discussion of extrasolar planets and the methods of their detection, with more than two hundred illustrations, mostly in color.

Deeg, Hans, Juan Antonio Belmonte, and Antonio Aparicio, eds. *Extrasolar Planets*. New York: Cambridge University Press, 2008. Eleven contributors to this book provide detailed and comprehensive discussions of detection, formation, statistical properties, and habitability of extrasolar planets.

Dvorak, Rudolf, ed. *Extrasolar Planets: Formation, Detection, and Dynamics*. Weinheim, Germany: Wiley-VCH, 2008. The book comprises eleven papers reviewing research on the detection and analysis research of extrasolar planets and their properties.

Goldsmith, Donald. *Worlds Unnumbered: The Search for Extrasolar Planets*. Illustrations by Jon Lomberg. Sausalito, Calif.: University Science Books, 2002. This introduction to extrasolar planet discoveries and their significance is written by an experienced and authoritative science writer.

Miller, Ron. *Extrasolar Planets*. Minneapolis, Minn.: Twenty-First Century Books, 2002. This book, written for middle school students, includes historical background in astronomy leading up to a description of extrasolar planet discoveries and is colorfully illustrated, including some planetary landscapes by the author.

See also: Extrasolar Planets; Extraterrestrial Life in the Solar System; Infrared Astronomy; Jupiter's Interior; Planetary Formation; Protostars; Solar System: Origins.

Extraterrestrial Life in the Solar System

Category: Life in the Solar System

Exobiology is the search for and study of life on celestial bodies other than Earth. Within the solar system, Mars, Jupiter's satellite Europa, and Saturn's satellites Titan and Enceladus are all considered possible sites where life, or the precursor chemistry needed for the rise of primitive living organisms, might have developed.

OVERVIEW

Understanding where life might have developed in the solar system requires comprehension of how life arose on Earth. The earliest evidence of life on Earth is the presence of organic matter derived from biological processes recorded in rocks that are about 3.2 billion years old. Life may have developed very early in Earth's history. However, much of the fossil record of early life on the Earth has been erased by subsequent geophysical activity. Biologists have pieced together some of that early history by examining the remaining fossil record and by performing a series of laboratory experiments.

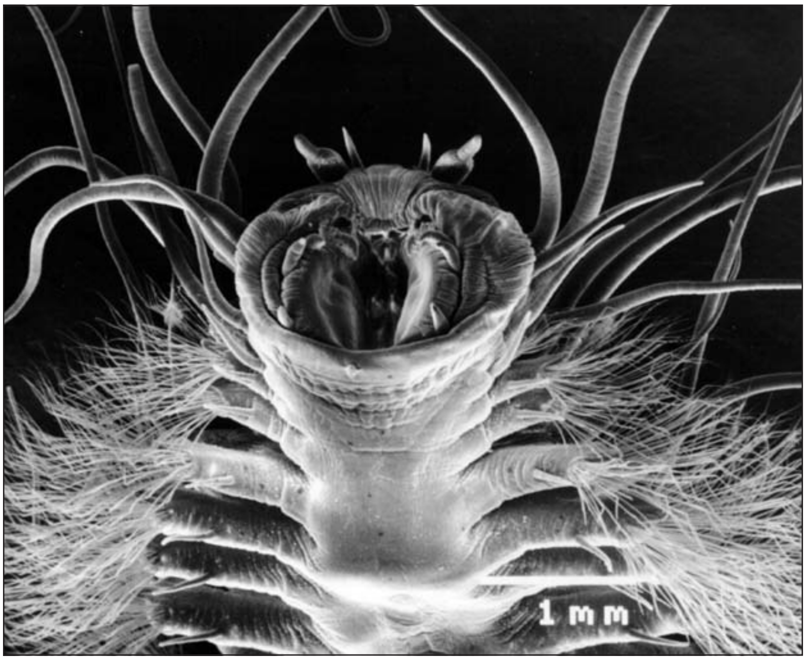
Life on Earth is based on complex organic molecules, consisting of chains of carbon, hydrogen, nitrogen, and oxygen. However, organic molecules can be produced by simple chemical reactions as well as by biological activity. Thus, to determine if a process is truly biological, rather than simply a chemical reaction, it is necessary to define the criteria for life. The ability of an organism to reproduce itself is considered to be an essential feature of life. Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are the organic molecules that control heredity in terrestrial life-forms. Thus, DNA and RNA are considered essential for reproduction of life on Earth. These two nucleic acids are produced only with the help of certain proteins. A major focus of exobiology is to understand how DNA, RNA, and the proteins essential in their production originated.

A major breakthrough occurred in 1953, when Stanley Miller, a graduate student at the

University of Chicago, and his research supervisor, Professor Harold Urey, produced amino acids, the basic building blocks of proteins, in a sealed environment simulating conditions believed to be present on the early Earth. Miller and Urey continuously passed electrical sparks through a chamber filled with a gaseous mixture of methane, ammonia, and hydrogen (a composition believed to be similar to that of the early atmosphere of the Earth) and water vapor (representing the water contributed by the Earth's oceans). After several days they extracted a mixture of organic molecules, including amino acids, from the bottom of the chamber. The Miller-Urey experiment suggested that lightning discharges throughout Earth's early atmosphere could have deposited amino acids onto the planet's surface. Other experiments demonstrated that bombardment of the gas mixture by high-energy particles, simulating cosmic rays, produced similar results. These experiments suggest that three sufficient conditions must be met to produce amino acids: a supply of carbon-rich material must be present, liquid water must be available, and some energy source (electrical discharge, high-energy particles, or possibly heat and sunlight) is required. Although sufficient, it remains to be proven that these three conditions are also necessary.

Scientists have examined the planets and satellites of the solar system, searching for locations where all three conditions are met. Water may be the most critical restriction, since it remains a liquid only over a very narrow range of temperatures. The surfaces of Venus and Mercury are too hot for liquid water to be present. Jupiter, Saturn, Uranus, Neptune, and Pluto are too cold to support liquid water. Thus, of these nine planets only Earth and Mars seem to be suitable candidates for life, because they are in the range of distances from the Sun such that they could support liquid water. Venus is within the Sun's habitable zone as well, but a runaway greenhouse effect of unknown origin has left it inhospitable to life as we understand and recognize it.

Where life is abundant, it can produce changes in the atmosphere of a planet, allowing astronomers to search for unusual signatures of biological activity. The present composition of the Earth's atmosphere, dominated by nitrogen and oxygen, is regulated by the life-cycle processes of respiration and photosynthesis of Earthly organisms. The atmosphere of Mars, on the other hand, is dominated by carbon dioxide, and it contains only a trace amount of oxygen. Thus, by the 1960's astronomers had observed that, at least in the present era, living organisms were not present in sufficient abundance to perturb the atmospheric chemistry of Mars.



Some of Earth's strange life-forms tell us that it is possible for life to exist in extreme conditions. This 2-inch, centipede-like worm, for example, was found in 1997 living within mounds of methane on the floor of the Gulf of Mexico. (NOAA)

The beginning of the space age made it possible to employ robotic spacecraft to perform direct measurements on the surface of some planets in the expanding search for evidence of life. The first search was performed on Mars by two Viking spacecraft, developed by the National Aeronautics and Space Administration (NASA), which landed safely in 1976. Each Viking spacecraft was equipped with instruments designed to examine the soils of Mars for evidence of Earth-like life.

During the 1980's and 1990's, developments in terrestrial biology changed how exobiologists looked at the essential conditions for the development of life. Single-celled organisms called archaebacteria, which may have developed very early in Earth's history, were discovered. These archaebacteria live in oxygen-deprived places, such as the hot springs at Yellowstone National Park. Archaebacteria take in carbon dioxide and give off methane, and they actually cannot thrive in the presence of oxygen. They have genetic material different from that of other terrestrial life-forms, suggesting that they possibly evolved independently from the more common life-forms very early in Earth's history at a time before the current oxygen-rich atmosphere arose. Other terrestrial microorganisms were discovered that live on sulfur from geothermal sources rather than by relying on the Sun to supply energy. The discovery of these unusual terrestrial life-forms suggests that conditions required for development of the common forms of life on Earth may not be required for the development of all life. Thus, some planets and/or their satellites previously believed to be unsuitable for the development of life may be habitable by organisms rather different from the common life-forms on Earth. This complicates the search for extraterrestrial life, because many experiments, such as those conducted by the Viking spacecraft on Mars in 1976, look only for signatures specific to common terrestrial life.

In the late 1990's and throughout the first decade of the twenty-first century, extrasolar planets were increasingly detected. Although the majority of the first hundred or so worlds were hot Jupiters or at least bizarre large planets in systems not conducive to life as we understand it, in time it became increasingly clear that technology would shortly be capable of picking up Earth-sized planets. A space-based observatory named Kepler was readied for launch in late 2008. One of its planned objectives was to expand the list of extrasolar planets tremendously and perhaps detect the first Earth-like planets.

Three things must be noted, however. One, there are scientists who dispute the Miller-Urey experiment's validity in terms of the suggestion that production of organic materials in this fashion necessarily leads to the development of life. Two, there remain—even three decades after the Viking biological experiment produced data suggestive of superoxide reactions in the Martian soil rather than a biological metabolism—many scientists who believe that the Viking results were misinterpreted. Perhaps the dismissal of a biological result was premature. Third, many astrobiologists insist that if one restricts one's search to life as we know it based on DNA, one severely limits the possibility for a successful result. Some even question the need for water as a "universal" solvent for life. Organisms using a different solvent would be vastly different.

METHODS OF STUDY

One focus of the search for life is to identify the carbon-rich compounds available for life's development. Impacts of meteorites, asteroids, and comets are believed to have contributed a carbon-rich layer to the Earth's early surface and other planets and their satellites. One particularly carbon-rich meteorite, called Murchison, fell in Australia in 1969. Detailed studies of Murchison established that it contains numerous organic compounds, including amino acids.

In 1986 five spacecraft, two launched by the Soviet Union, two by Japan, and one by the European Space Agency, flew past Halley's comet. Dust analyzers on some of these spacecraft determined the chemical composition of individual dust particles emitted by the comet. These instruments detected a large number of carbon-rich particles, many of which also contained hydrogen, suggesting the presence of organic molecules in the dust. However, detailed analysis of organic molecules requires sophisticated scientific instruments too large and complicated to be flown on those spacecraft. NASA launched a spacecraft called Stardust to fly to Comet Wild 2 to collect dust emitted by that comet. It successfully returned samples to Earth in 2006. Laboratory study of the dust established the abundances and types of organic compounds present in Wild 2.

The second focus of the search for life is to perform direct tests for the presence of biological activity on other planets or satellites. Apollo astronauts collected the first samples from the Moon in 1969. When they returned to Earth, the astronauts, their spacecraft, and their prized lunar rocks were subjected to a twenty-one-day quarantine during which scientists searched for living microorganisms that might be hazardous to life on Earth. Fragments from lunar rocks were crushed and placed in a standard culture medium, a nutrient-rich soup that promotes the growth of microorganisms. Microscopic examination of these samples showed no evidence of living microorganisms. More detailed studies of the lunar rocks have shown no fossil evidence of life-forms that might once have developed on the Moon but are now extinct. Examination of lunar samples revealed them to be exceptionally dry, with none showing any evidence of liquid water. The absence of liquid water was taken to indicate that the Moon was always a lifeless body.

Initial experiments in the search for life on another planet were conducted in 1976 by the two Viking spacecraft that landed on Mars. Each Viking carried four instruments to examine the soil samples for evidence of such basic life-cycle processes as respiration or photosynthesis. The Gas Exchange Experiment deposited samples of Martian soil in a chamber containing a culture medium. This apparatus monitored the composition of gas within its chamber, looking for changes in the abundance of carbon dioxide, oxygen, or hydrogen that would signal metabolic activity by microorganisms in the soil.

In a second experiment, the Labeled Release Experiment, radioactive carbon atoms were incorporated into the culture medium. A detector looked for the appearance of radioactive carbon in released gas, signaling that the addition of Martian soil to the nutrient had resulted in a reaction of biological origin. Both experiments produced positive results, but the effects were much more dramatic than the scientists had expected. These positive results were eventually explained as chemical reactions initiated because of the highly reactive nature of the surface materials on Mars resulting from their exposure to ultraviolet light from the Sun, a superoxide chemical reaction.

The Pyrolitic Release Experiment provided an opportunity to test that explanation. It was also a labeled release experiment, but this apparatus had the additional capability of heating soil samples between experiments. Scientists heated soil to 548 kelvins, well above the temperature expected to kill any microorganisms present in the soil. Even then the Pyrolitic Release Experiment yielded positive results, suggesting that the release was produced by a chemical reaction involving superoxides rather than a biological process.

A fourth experiment, the Gas Chromatograph Mass Spectrometry Experiment, produced the most convincing evidence that the soils at the Viking landing sites contained no microorganisms. This instrument found no organic molecules within the soil down to a limit of a few parts per million. Even the organic molecules that would be expected in the soils from the accumulation of meteorites like Murchison were absent. Subsequent studies indicated that high chemical reactivity of the soils as well as intense ultraviolet radiation striking the surface would rapidly destroy most organic molecules. Thus, if there is life on Mars, the two Viking spacecraft, which were able only to sample the near-surface soils, were probably looking in the wrong places.

Although instruments on both Viking landers found no evidence of biological activity in their soil samples, the two Viking orbiters obtained high-resolution photographs of Mars's surface, producing results which excited exobiologists. Several regions on Mars revealed features similar to extensive water flow channels on Earth, leading many geologists to conclude that water had flowed freely on the surface of Mars at some earlier period in its history. Because of the assumed importance of liquid water in the development of life, some exobiologists suggested that life might have developed on Mars in that earlier era and that life might now exist in subsurface layers protected from ultraviolet radiation. Or perhaps such life had gone extinct, leaving only fossil evidence behind.

In 1996 scientists from NASA's Johnson Space Center reported that a meteorite called ALH 84001, one ejected from the surface of Mars and deposited in the Antarctic about thirteen thousand years ago, contained microscopic features that might indicate ancient Martian biological activity. This resulted in renewed interest in the search for life on Mars. These suspected fossils resembled wormlike creatures but their size was extraordinarily small. Many scientists pointed out that similar nanometer-sized structures could be produced geochemically and had nothing to do with life. This dispute has not yet been resolved.

After the 1997 Mars Pathfinder exploration of the Red Planet returned amazing images of rocks and terrain, NASA planned a series of robotic spacecraft to continue the exploration of Mars. Two of those spacecraft, the Mars Exploration Rovers named Spirit and Opportunity, launched in June and July, 2003, respectively. They successfully landed on Mars in early 2004 and spent at least the next four years moving about their landing sites searching for evidence of water. Later spacecraft were intended to be even more ambitious, leading to the ultimate desire of exobiologists and planetary scientists alike: a sample return mission from Mars sometime in the second decade of the twenty-first century.

The same techniques used to search for current or fossil life on Mars can be applied to other planets or satellites that are identified as suitable candidates for the development of life. The Galileo spacecraft, placed in orbit around Jupiter in late 1995, obtained close-up photographs of Jupiter's four largest satellites. One of these, Europa, emerged as another potential site for the development of life. One of Galileo's orbits around Jupiter took it within 363 miles of Europa's surface, allowing its cameras to photograph objects as small as 75 feet across. These images showed evidence of ice flows that had broken from a solid sheet and been displaced, suggesting that they had floated or slipped across a liquid ocean or on a layer of slush below. Calculations indicated that Jupiter's extreme gravitational pull could introduce tidal distortions that produce sufficient heat to allow liquid water to exist beneath Europa's icy surface. Other photographs showed dark deposits, possibly carbon-rich material contributed by meteorites.

Titan, the largest satellite of Saturn, has a methane-rich atmosphere believed to be similar in composition to that of the early Earth. High-energy electrons and protons, trapped in the magnetic field of Saturn, continually bombard the upper region of Titan's atmosphere. This bombardment is believed to produce complex organic molecules that rain down onto Titan's surface. Titan is too cold to have liquid water. Titan remained the primary target of study for the Cassini spacecraft, which was launched in October, 1997, and arrived in the Saturn system in early July, 2004. Cassini dropped its Huygens probe, loaded with instruments to measure the types and abundances of the organic molecules, into Titan's atmosphere. The Huygens probe showed its surface may be covered with lakes of methane or ethane, which some scientists now speculate might be sufficient to allow primitive life to develop. Also, Titan's crust appears to move significantly as if floating on a subsurface ocean, adding another intriguing aspect to the possibility of organic chemistry and/or primitive life on Titan.

Even Enceladus displays unexpected geyser activity at its south polar regions. This suggested the possibility of liquid water underneath the surface and therefore the potential for primitive life. Neptune's Triton also exhibits cryovolcanism at an even lower temperature. More research is needed to determine the nature of this mechanism, and that investigation would likely have to await a Neptune orbiter.

Exobiologists were excited to see the possible existence of the three conditions believed necessary for the development of life: carbon-rich material, water, and energy from the Jovian tides. Several follow-on missions have been suggested. A spacecraft placed into orbit around Europa could use radar to see through several miles of ice, detecting any water below and providing a clear test of the ocean model. More ambitious proposals include a spacecraft that would fling a 9-kilogram projectile into the surface of Europa, catch some of the debris lofted by the collision, and return it to terrestrial laboratories for examination. Another common proposal would see a submersible vehicle melt its way through Europa's icy crust to reach a potential subsurface layer of liquid water and image the local environment directly.

CONTEXT

The possibility that life might have developed elsewhere in the solar system has been the subject of speculation for hundreds of years. In 1820, Carl Gauss, a German mathematician, suggested cutting geometrical patterns into the Siberian forest large enough to be seen by an observer using a telescope from the Moon or Mars. The idea was to motivate any inhabitants of the Moon or Mars to engineer similar geometrical patterns, initiating crude communication with the Earth. Other suggestions for communication with intelligent life included setting huge fires in the Sahara desert and constructing large mirrors to reflect sunlight back into space. These early ideas of how to communicate with intelligent life elsewhere in the solar system did not focus on particular sites where the conditions were expected to be appropriate for the development of life.

Although its origins go back as far as 1929, radio astronomy only gained respect within the astronomical community in the early 1950's in the aftermath of World War II, when radio equipment necessary to "listen" to the heavens became available as war surplus. Radio astronomers soon discovered that the natural universe was far from radio quiet. Some scientists, beginning with astronomer Frank Drake, wondered about and then tested the idea that intelligences beyond Earth might be transmitting recognizable radio signals. In due time a coordinated Search for Extraterrestrial Intelligence (SETI) program was developed. No verifiable signals of intelligence have yet been received from deep space.

Only in the second half of the twentieth century did biologists begin to develop an understanding of how life originated on Earth. This knowledge provided clues as to the conditions needed for similar forms of life to develop elsewhere in the solar system. The study of terrestrial life indicates that it originated as simple, single-celled microorganisms and that these simple microorganisms might develop quickly and easily on other planets and/or their satellites as well. Thus, the focus of solar system exobiology shifted from the search for intelligent life, which has not been seen on any planet other than Earth, to the search for simple microorganisms. However, SETI continued, although for a time Congress removed any support for the project through NASA's federal allocations. In time commercial funds supplemented federal funding for SETI projects. For a time in the period following release of the popular movie *Contact* (1997), based on a book by the late Carl Sagan, public interest in SETI increased dramatically.

The dawn of the space age inaugurated an era when spacecraft could be used to search for environments favorable to the development of life, perform experiments designed to detect living organisms on the surface of other planets and/or their satellites, and ultimately return samples to Earth so that scientists could examine them for evidence of biological activity or fossil evidence of past life. Although scientific interest in life elsewhere in the solar system reached a low point after the negative results of the Viking landers in 1976, there was a resurgence of interest by the end of the twentieth century. Discovery of river channels on Mars, possible fossil evidence for ancient microorganisms in a meteorite from Mars, hints of water ice on the Moon and Mercury, oceans on Europa and Enceladus, organic materials and an atmosphere on Titan, and cryovolcanism seen on Enceladus and Triton suggest that the solar system might not be as inhospitable to the development of life as was believed immediately following the results of the Viking landers.

George J. Flynn

FURTHER READING

Goldsmith, Donald, and Tobias Owen. *The Search for Life in the Universe*. 3d ed. New York: University Science Books, 2001. Speculates scientifically about the possibility of the existence of intelligent beings beyond Earth.

Greenberg, Richard. *Europa the Ocean Moon: Search for an Alien Biosphere*. Berlin: Springer, 2005. A complete description of current knowledge of Europa through the post-Galileo spacecraft era. Discussion of the astrobiological implications of an ocean underneath Europa's icy crust. Well illustrated and accessible to astronomy enthusiasts and college undergraduates.

Hansson, Anders. *Mars and the Development of Life*. New York: Ellis Horwood, 1991. A comprehensive, well-illustrated discussion of the conditions for the development of life and the search for life on Mars.

Harland, David M. *Cassini at Saturn: Huygens Results*. New York: Springer, 2007. Provides a thorough explanation of the entire Cassini program, including the Huygens landing on Saturn's largest satellite. Essentially a complete collection of NASA releases from the start of Cassini flight operations through the majority of Cassini's seventy orbits during its primary mission. Cassini's primary mission concluded a year after this book was published. Technical writing style, but accessible to a wide audience.

_______. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002. A technical description of the Cassini program, its science goals, and the instruments used to accomplish those goals. Written before Cassini arrived at Saturn. Provides a historical review of pre-Cassini knowledge of the Saturn system.

Michaud, Michael A. G. *Contact with Alien Civilizations: Our Hopes and Fears About Encountering Extraterrestrials*. New York: Springer, 2006. Explores the possibility of extraterrestrial intelligence, speculates about human-extraterrestrial interactions, and discusses the impact on society that making contact could have.

Orgel, Leslie. "The Origin of Life on Earth." *Scientific American* 271 (October, 1994): 76-83. A comprehensive description of how the emergence of RNA is believed to have been critical to the development of life on Earth. Includes a good account of the Miller-Urey synthesis experiment.

Sagan, Carl. *Contact*. New York: Pocket, 1997. This novel by the Cornell astronomer and science popularizer provides an account of contact with a greatly advanced intelligent species beyond Earth. With all the real science that is included, the novel easily allows the reader to suspend belief in those areas where the science is highly speculative. The book prompted the production of a popular movie by the same name.

_______. "The Search for Extraterrestrial Life." *Scientific American* 271 (October, 1994): 92-99. A clearly written, well-illustrated account focusing on the scientific results of the Viking spacecraft and plans to investigate the atmosphere of Titan using the Cassini-Huygens spacecraft.

Squyres, Steve. *Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet*. New York: Hyperion, 2006. Written by the principal investigator for the Mars Exploration Rovers Spirit and Opportunity, this fascinating book provides a general audience with a behind-the-scenes look at how robotic missions to the planets are planned, funded, developed, and flown. A personal story of excitement, frustrations, the satisfaction of overcoming difficulties, and the ongoing thrills of discovery.

Ward, Peter. *Life as We Do Not Know It: The NASA Search for (and Synthesis of) Alien Life*. New York: Penguin, 2007. Written by a paleontologist, this book presents reasonable speculations about the nature of primitive life that may someday be found through space exploration or extrapolation from remote-sensing data.

Webb, Stephen. *If the Universe Is Teeming with Aliens . . . Where Is Everybody? Fifty Solutions to Fermi's Paradox and the Problem of Extraterrestrial Life*. New York: Springer, 2002. The great physicist Enrico Fermi declared that if intelligences greater than those of humans existed, then they would already be here visiting on Earth. This book looks at various solutions to surmounting the vast distances between star systems and the nature of extraterrestrial life in the context of answering Fermi's paradox.

See also: Habitable Zones; Life's Origins; Main Sequence Stars; Mars: Possible Life; Search for Extraterrestrial Intelligence.

G

Gamma-Ray Bursters

Category: The Stellar Context

Gamma-ray bursts for a long time were an unexplained phenomenon in high-energy astrophysics. A variety of spacecraft have detected and studied these random, brief, and intense bursts of gamma rays, which come from all parts of the sky. Most theories associate them with neutron stars in the Milky Way galaxy, but an extragalactic source cannot be excluded.

OVERVIEW

Gamma-ray bursts (GRBs) constitute a unique phenomenon in astronomy. During their brief appearance, they are brighter than all other objects in the sky, including the Sun. About a hundred strong GRBs occur every year. It is not known how many weaker bursts occur. For many years after their discovery, their source represented one of the greatest mysteries in astrophysics. They occur at random times and appear to be randomly distributed over the sky. There is no particular clustering of GRBs in any region, and until the Swift spacecraft was launched in 2004, they had not been associated with any known objects. GRBs are extremely difficult to study, since it is never known when or where a GRB will occur.

It has become recognized that there are distinct classes of GRBs, with different properties, that may be caused by entirely different objects or emission mechanisms. The situation is analogous to the recognition, long after the first telescopes came into use, that not all nonstellar objects should be classified simply as "nebulae," since they include objects as diverse as supernova remnants, galaxies, star clusters, and planetary nebulae. It is believed that, similarly, the phenomena that scientists call "gamma-ray bursts" will be found to be caused by more than one process or object.

Discovery of GRBs in 1972 by the Vela spacecraft was a classic case of a serendipitous discovery, a discovery made while looking for something else. The Vela spacecraft were designed and operated to detect nuclear explosions from space. This series of small spacecraft, built by TRW and launched in the mid-1960's, contained a wide variety of sensors that "looked" in all directions. The spacecraft were launched into high, eccentric orbits so that they could even scan the area behind the Moon for clandestine nuclear explosions.

Gamma-ray detectors aboard the Vela spacecraft were designed and built at the Los Alamos National Laboratory. They consisted of small scintillation detectors, the output of which was continuously monitored for an increased rate above background. After several years of operation, occasional triggers were detected, but they were dismissed since no other sensors on board the spacecraft recorded the events and because such "glitches" were common to detectors on other spacecraft. It was not until Los Alamos scientists began studying these triggers in greater detail that their nature became known. In many cases, two or more spacecraft would record a trigger at nearly the same time. It was first suspected that a source of gamma rays from Earth, the Sun, or another object or region within the solar system was causing the GRBs. When precise gamma-ray-burst timing analysis was performed, it became evident that triggers were caused by a plane wave of gamma rays striking the array of widely separated spacecraft. This type of wave could be caused only by a powerful point source of gamma rays far beyond the solar system.

Los Alamos scientists announced their discovery at a meeting of the American Astronomical Society in Columbus, Ohio, in 1973 and published their findings in an astrophysical journal. Almost immediately, there was a flurry of activity to try to explain GRBs and to obtain more experimental data. As experimenters began to look through old data and data from still-operating spacecraft, many confirmed GRBs were uncovered in addition to those detected by the Vela spacecraft. Among the earlier spacecraft that confirmed the existence of GRBs were the Orbiting Solar Observatories, the Orbiting Geophysical Observatories, the Small Astronomy Satellites, the Interplanetary Monitoring Platform, Kosmos 461, Apollo 16, and the German spacecraft TD 1. It should be noted that none of these spacecraft had detectors that were specifically designed to detect GRBs. It was only because of the intensity of the bursts and their coincidence with other observations that they were detectable by instruments designed for other purposes.

By the late 1970's, a network of small detectors on interplanetary spacecraft was established in an attempt to locate the source of the GRBs more precisely. Included in this network were instruments aboard the Pioneer Venus orbiter; International Sun-Earth Explorer 3 (ISEE 3); Veneras 11, 12, 13, and 14; Prognoz 7; and Helios 2. For the first time, these spacecraft provided long interplanetary baseline distances required to locate the GRBs within one arc minute. Unfortunately, with one important exception, no unusual objects were detected near the burst sources. The exception was the GRB of March 5, 1979. It occurred in or near a supernova remnant in the Large Magellanic Cloud. This burst was unusual in other respects, however, so it may have been part of a separate class of GRBs.

The Russians, in collaboration with Bulgaria, Denmark, and France, launched the Granat observatory late in 1989. Outfitted with two instruments to investigate high-energy astrophysical objects from X-ray to gamma-ray wavelengths, Granat lasted until 1998. However, five years into its orbital lifetime the observatory ran out of attitude control gas, making directional surveys no longer possible. Among Granat's greatest discoveries were the detection of electron-positron annihilation from a galactic microquasar, nineteen GRBs, and the identification of numerous objects that were candidates for black holes.

The Gamma Ray Observatory (GRO) was deployed into an independent orbit from the space shuttle Atlantis during the STS-37 mission in early April, 1991. The second in the National Aeronautics and Space Administration's (NASA's) Great Observatory series, once in orbit GRO was renamed the Compton Gamma Ray Observatory (CGRO or just Compton Observatory) after the Nobel Prize-winning physicist Arthur Compton, for whom an important effect involving an interaction of matter and electromagnetic radiation is also named. That interaction involves a shift in wavelength when a photon is "scattered" off a free electron. This effect is most pronounced in X-ray and gamma-ray photons.

CGRO was left in a low, nearly circular Earth orbit (at an altitude of 450 kilometers) to keep it out of the Van Allen radiation belts. Ideally it would have been deployed at an altitude well above those radiation zones, but it was far too heavy (17,000 kilograms) for the shuttle to put it up that high; use of an upper stage in concert with the shuttle carrying such a heavy payload was out of the realm of possibility as well.

Compton's most important improvement over previous gamma ray-detecting satellites was that its suite of four experiments covered energies ranging from as low as 20 kilo-electron volts (keV) to 30 giga-electron volts (GeV). No previous orbital detector had spanned six orders of magnitude in wavelength (or energy as a consequence) in the gamma-ray portion of the electromagnetic spectrum.

CGRO's four instruments were the Burst and Transient Source Experiment (BATSE), the Oriented Scintillation Spectrometer Experiment (OSSE), the Imaging Compton Telescope (COMPTEL), and the Energetic Gamma Ray Experiment Telescope (EGRET). Provided by NASA's Marshall Space Flight Center, BATSE was designed to search for short GRBs ranging in energy from 20 to 600 keV and also generate full-sky surveys for long-duration gamma-ray sources. Provided by the Naval Research Laboratory, OSSE was outfitted with four individually pointing detectors capable of picking up radiation from 0.05 to 10 mega-electron volts (MeV). A pair of these detectors would record emission from a source while the other two would record the background near those sources for contrast. COMPTEL was provided by a collaboration of the University of New Hampshire, the Netherlands Institute for

Space Research, the Max Planck Institute, and the European Space Agency's (ESA's) Astrophysics Division. COMPTEL was designed with a wide field of view and sensors capable of identifying sources (in an energy range of 0.75 to 30 MeV) to within one degree. EGRET was provided by a collaboration of NASA's Goddard Space Flight Center, Stanford University, and the Max Planck Institute. This was the highest energy detecting portion of CGRO, capable of recording emissions in the range of 20 MeV to 30 GeV and identifying the location of incoming radiation to within a fraction of a degree; that level of resolution was useful in having other satellites precisely locate sources picked up by EGRET.

Compton was not capable of being repaired in orbit like the Hubble Space Telescope. CGRO's systems began to degrade, especially threatening the loss of attitude control. It was therefore decided to drive the observatory into the atmosphere over a portion of the Pacific Ocean that was sparsely populated at best. Large debris that would survive reentry, such as a major portion of one of Compton's detectors, would drop harmlessly into the ocean. CGRO was deorbited on June 4, 2000, reluctantly ending nine years of unprecedented gamma-ray astrophysics research.

The Swift Gamma Ray Burst spacecraft was launched by a Delta II booster on November 20, 2004, and placed in an orbit 600 kilometers above the Earth's surface. Swift was designed to provide the best all-sky survey of gamma rays yet, to provide alerts to transient astrophysical events such as supernovae and GRBs of both short and long duration, to help identify the location of GRBs, and to assist in determining the distances to gamma-ray bursters at cosmological distances representing a time early in the universe's evolution. Swift was outfitted with just three instruments: the Burst Alert Telescope (BAT), X-ray Telescope (XRT), and UltraViolet/Optical Telescope (UVDT). The most important aspect of Swift's capability was the rapidity with which it could respond to a gamma-ray detector and precisely locate its source so that these three instruments and other assets available worldwide to the astronomy community could record and study the afterglow of a GRB. Within about fifteen seconds on average, Swift could identify the source of gamma rays to within approximately one arc minute of the sky. The Swift Mission Operation Center, located on the campus of Pennsylvania State University, serves as a clearinghouse alerting other astronomers around the world to GRB events.

NASA's Gamma-Ray Large Area Space Telescope (GLAST) was launched by a Delta II booster in June 11, 2008. GLAST involved an international team consisting of space agencies and research groups from the United States, France, Germany, Italy, Sweden, and Japan. GLAST was designed as a follow-on gamma-ray astronomical observatory to the lost Compton GRO; however, it was not considered a member of NASA's Great Observatory program. It was intended to investigate cosmological questions raised by Compton observations as well as energies far in excess of what can be produced in particle accelerators on Earth. Scientists expected to use GLAST to obtain better understanding of black holes, neutron stars, and high-speed gas and how they produce gamma radiation.

KNOWLEDGE GAINED

Since GRBs have not been identified with known objects, their distance is highly uncertain. This, in turn, makes it difficult to speculate on their origin. Since the distance to the burst sources is not known, the intrinsic luminosity of the source is even more uncertain. Many of the early theories of GRBs posited exotic phenomena or objects to explain them. In later years, most models have associated GRBs with explosive events near, or at the surfaces of, neutron stars within the Milky Way galaxy. These explosions could be caused by thermonuclear reactions resulting from the collision of interstellar material, comets, or asteroids with neutron stars, or from the annihilation of strong magnetic fields near such stars. Another theory attributes GRBs to a sudden shift of the solid crust that is thought to be present in neutron stars. There are also models of GRBs that attribute them to enormous explosions occurring at cosmological distances, or distances near the edge of the observable universe. At these distances, the luminosity of a GRB would be equivalent to that of a supernova, although all of its energy would be emitted at gamma-ray wavelengths and within the duration of a GRB.

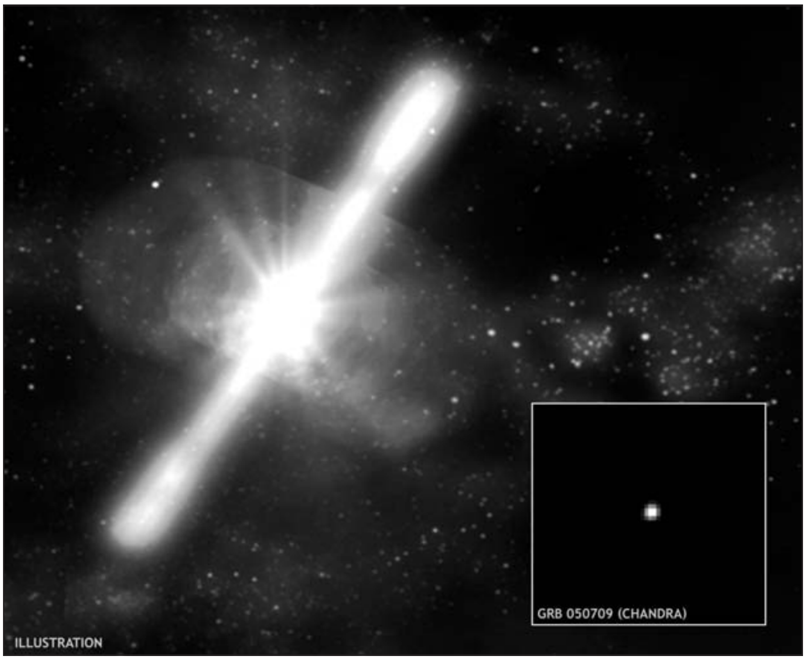
The three observable properties of GRBs that are most often studied are their time histories, their energy spectra, and the statistical properties of their intensity and distribution over the sky. Attempts to locate a GRB and identify it with a known object have thus far been unsuccessful. Very sensitive optical, radio, and X-ray searches have been made of precisely located gamma-ray-burst "error boxes" (the region of uncertainty in the position of a celestial source). These searches have been either inconclusive or controversial. A search of old photographic plates from telescopes in the Southern Hemisphere, however, has shown two or three transient, starlike optical images at the locations of GRBs. The authenticity and the significance of these observations are still being debated.

The time history of a GRB refers to the intensity variations of the burst as a function of time. Some GRBs show extremely rapid fluctuations over their entire duration, which may encompass a minute or two. Other bursts last a few seconds, during which time only smooth variations are seen. Still others exhibit a single spike lasting only a fraction of a second. The rapid variations indicate that the source of a GRB is a very small region or compact object, such as a neutron star or a black hole. The GRB of March 5, 1979, was unique in that it had a single, intense spike that was followed by a lower-level emission with an eight-second period that lasted for more than two hundred seconds.

The spectra of GRBs indicate that the sources contain regions of extremely high temperatures—perhaps the highest in the universe. In many cases, the gamma-ray energies extend up to 100 MeV. Extremely rapid variations are observed in the spectra of most GRBs. In addition, gamma-ray line features are observed that may be explained by the gamma rays coming from regions with extremely high magnetic fields, such as those expected near neutron stars. There is also a class of GRBs that have softer spectra, or lower temperatures, and are observed to be repetitive. Soviet researchers also reported gamma-ray line features near 400 keV, which some interpreted as caused by the annihilation and redshift of electron-positron plasma near the surface of a neutron star.

Granat began working in a survey mode in late 1994. Before the demise of this Russian observatory, data from Granat detailed spectral and temporal variability of potential black holes, discovered the specific radiation emerging from electron-positron annihilation from the X-ray nova Muscae and a galactic microquasar, and improved imaging of the Milky Way's galactic center.

In February, 1997, telescopes sighted the source of a burst, a diffuse, elongated object with a bright core. Astronomers thought the ob-



This artist's rendering of GRB 050709 depicts a gamma-ray burst discovered July 9, 2005, by NASA's High-Energy Transient Explorer. The inset shows the X-ray afterglow visible three days later. (NASA/CXC/Caltech/D. Fox et al./ Illustration: NASA/D. Berry) ject might be a very distant galaxy, but they could not be certain without more data.

Compton provided a means whereby GRBs of short duration could be differentiated from longer-duration bursts. Its instruments picked up the first known gamma-ray repeaters. Over its lifetime, BATSE picked up at least one event per day, giving it a total of 2,700 gamma-ray detections. OSSE was used to survey the Milky Way's center and provided evidence suggesting that a cloud of antimatter existed exterior to the central region. COMPTEL observed gamma rays originating from the decay of the radioactive isotope aluminum 26 (^{26}Al) and generated a full-sky survey of that emission. EGRET's full-sky survey of high-energy gamma-ray emissions picked up 271 sources. However, only 100 of them were definitively identified. One surprise was the detection of terrestrial gamma rays originating from thunderstorms. CGRO instruments surveyed both pulsars and supernova remnants.

CGRO provided an enormous amount of data, and its instruments were still working at the time of its demise. Its data helped astrophysicists better characterize the nature of GRBs and provided the insight needed to design even better observations to follow, such as the Swift spacecraft and GLAST telescope.

The Swift observatory represented a major advance in the study of GRBs. Outfitted with detectors capable of observing from the visible through gamma-ray wavelengths, Swift was designed to scan the sky continuously for the signature of a GRBer. It had the capability to slew quickly and locate the source of the burst. At this point, the worldwide astronomical community would be alerted to use space-based and ground-based observatories to quickly measure and record the burster's afterglow. Although there were some anomalies with Swift's XRT instrument, the observatory was commissioned on February 1, 2005. It had detected an initial burster earlier on January 17, 2005, and was quickly able to identify the bright source in its field of view, thereby triggering an alert for further study by other observatories, such as the Chandra X-Ray Observatory. During its first year of operation alone, Swift found 90 GRBs. Swift was the first to detect the location of a short-duration GRB. GRB 050509b, observed by Swift on May 9, 2005, had a burst duration of merely 50 milliseconds; this demonstrated the rapid response time of Swift.

On September 4, 2005, Swift discovered the most distant known GRB, GRB 050904, located 12.6 billion light-years distant. In 2006, Swift identified the location of GRB 060614 to be 1.6 billion light-years distant. This burster lasted 102 seconds and had a signature indicating it most likely resulted from the formation of a black hole. Another great discovery came on January 9, 2008. While Swift was studying a supernova in NGC 2770, it detected an X-ray burst in the very same galaxy. This triggered coordinated studies with Chandra, the Very Large Array, the Keck I telescope, the 200- and 60-inch telescopes at Palomar, and the Hubble Space Telescope, making this perhaps the most intensely studied supernova at a very early point in its development, a study that was performed with instruments across the electromagnetic spectrum. On March 19, 2008, Swift detected four GRBs, a record for one day's observations. Even more important, the second of the four discovered that day, GRB 080319B, turned out to be the brightest celestial object ever detected. Located 7.5 billion light-years from Earth, this GRB was 2.5 million times brighter than any other supernova.

GLAST only achieved first light by the time this article was composed; thus little in the way of results could be provided. However, among the scientific objectives of GLAST research were the examination of high-energy astrophysical objects at energies greater than could be duplicated on Earth in laboratories; the search for sources of dark matter to illuminate the identity and physics of such exotic matter; investigation of gamma-ray bursters; investigation of the mechanisms whereby black holes produce jets and accelerate the material in them to speeds very close to that of light; and investigation of solar flares, cosmic rays, and pulsars at high energy.

CONTEXT

The field of high-energy astrophysics is a product of the space age. It is necessary to carry instruments and telescopes above Earth's atmosphere in order to observe the universe at X-ray and gamma-ray wavelengths. This relatively new branch of astronomy not only has taught scientists more about objects that they already knew to exist but also has revealed new types of objects and phenomena, including X-ray stars, black holes, and GRBs. These objects are among the most energetic and violent in the universe. Most of them are associated with the final stages in the life cycles of massive stars.

GRBs once represented the greatest unsolved problem in high-energy astrophysics. Their distance and luminosity were unknown for a long time. They are difficult to study because of their random and transient nature. Although the initial discovery and studies of GRBs were made by groups in the United States, the Soviets, often with French collaborators, also obtained many of the gamma-ray-burst data. The establishment of an international gamma-ray-burst observation network, combining data from as many as nine spacecraft, became a model for international collaboration in space exploration. It is expected that the continued study of GRBs not only will help astrophysicists in their understanding of these objects but also will enable scientists to study conditions of extreme temperature, pressure, and density that are unavailable anywhere else.

Gerald J. Fishman

FURTHER READING

Arny, Thomas T. *Explorations: An Introduction to Astronomy*. 3d ed. New York: McGraw-Hill, 2003. A general astronomy text for the nonscientist. Includes an interactive CD-ROM and is accompanied by a Web site. Comets are covered.

Fabian, A. C., K. A. Pounds, and R. D. Blandford. *Frontiers of X-ray Astronomy*. Cambridge, England: Cambridge University Press, 2004. For the most serious astronomy reader or students of astrophysics. Covers contemporary research with space-based X-ray telescopes. Also relevant to the subject of gamma-ray astronomy.

Gregory, Stephen A. *Introductory Astronomy and Astrophysics*. 4th ed. San Francisco: Brooks/Cole, 1997. Suitable as a textbook for introductory college-level course or an advanced high school course in general astronomy. Covers all topics, from solar-system bodies to cosmology. Contains some errors and issues with mathematical presentations.

Hillier, Rodney. *Gamma-Ray Astronomy*. Oxford: Clarendon Press, 1984. This well-illustrated and well-referenced book provides a comprehensive, college-level overview of the objects and methods of gamma-ray astronomy.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find the standard treatise for typical introductory courses aimed at too low a level. Suitable for an audience with varied science and mathematical backgrounds. Covers all topics from solar-system objects to cosmology.

Katz, Johnathan I. *High Energy Astrophysics*. Reading, Mass.: Addison-Wesley, 1987. A well-organized book covering all aspects of high-energy astrophysics but concentrating on emission mechanisms. For advanced readers.

Maccarone, Thomas J. *From X-ray Binaries to Quasars: Black Holes on All Mass Scales*. New York: Kindle, 2006. Provides descriptions of high-energy processes that produce X-ray emissions. Describes the cosmological significance of quasars, black holes, and other high-energy objects.

Verschuur, Gerrit L. *The Invisible Universe: The Story of Radio Astronomy*. New York: Springer Praxis, 2006. Provides a history of developments in radio astronomy and along the way describes the discovery of pulsars, quasars, and radio galaxies. Suitable for a general college science course as well as for astronomy majors as background information.

Weeks, T. C. *Very High Energy Gamma Ray Astronomy*. New York: Taylor & Francis, 2003. Covers gamma-ray astronomy through results from the Compton Gamma Ray Observatory. For students of either theoretical or experimental high-energy astrophysics.

See also: Brown Dwarfs; Main Sequence Stars; Novae, Bursters, and X-ray Sources; Nuclear

Synthesis in Stars; Protostars; Pulsars; Red Dwarf Stars; Red Giant Stars; Stellar Evolution; Supernovae; Thermonuclear Reactions in Stars; White and Black Dwarfs.

Ganymede

Categories: The Jovian System; Natural Planetary Satellites

The Jovian moon Ganymede was the first natural satellite, other than Earth's moon, to be discovered. It is also the largest satellite in the solar system—large enough to generate its own magnetic field, an unusual characteristic for a satellite.

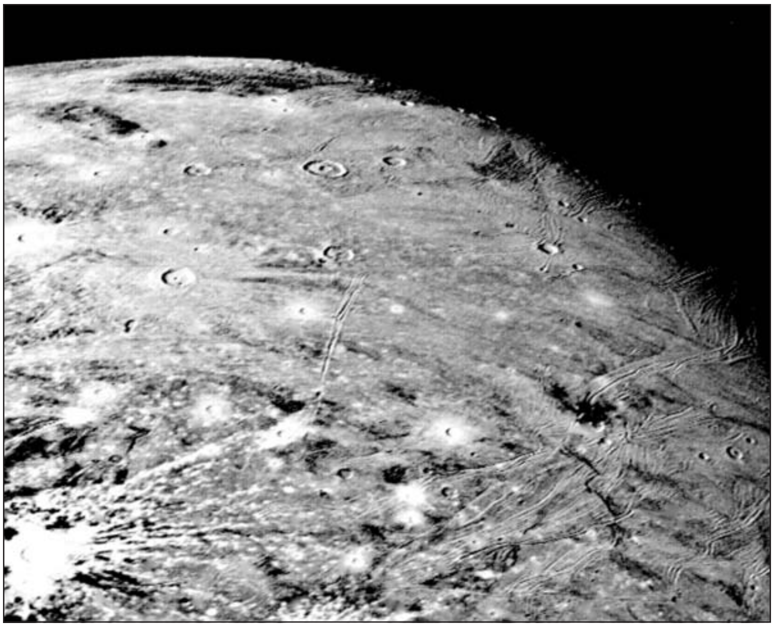
OVERVIEW

Ganymede, the largest satellite of Jupiter, was discovered by Galileo Galilei with a telescope in 1610. He published the information in *Siderius Nuncius* (starry messenger) and thereby initiated a dispute with the Church that would eventually lead him to be placed under house arrest for the remainder of his life; it was heresy to say that anything revolved around something other than the Earth. The satellite was named by Simon Marius for one of the lovers of the Roman god Jupiter.

Ganymede is 5,280 kilometers in diameter, and just over 1 million kilometers from Jupiter; it is the seventh of sixteen satellites. It is common for a satellite to present the same face to its planet at all times—a relationship called synchronous—and Ganymede does this, just as Earth's moon does. The rotation of Ganymede is prograde, that is, in the same direction as that of Jupiter. Ganymede's orbit is almost circular, meaning that its eccentricity (the measure of how close to a circular orbit the satellite travels) is small. A circular orbit has an eccentricity of zero. Ganymede's angle of inclination is less than a degree, meaning that this moon rotates almost exactly in the plane of Jupiter's equator.

Ganymede's albedo, the amount of sunlight reflected, is large. This reflectivity is caused by ice mixed with carbon-rich soil on the surface of the satellite. When the ice underneath the surface is heated and melts, it erupts to the surface. The soil, which is denser than water, sinks below the water. The water freezes, causing a bright spot on the surface. The water is heated either by radioactive decay or by tidal flexing. Not only does the gravity of Jupiter and Callisto pull on Ganymede; the moon also has Laplace resonance, which occurs because of the forces from the satellites Io and Europa. Every time Ganymede rotates around Jupiter once, Eu-

This image of Jupiter's moon Ganymede shows the seismic activity in the crust and outer coating of water ice that has led to this body's nickname, the Moonquake World. It may be that Ganymede, like Earth, has plate tectonic activity. (NASA, Voyager, © Calvin J. Hamilton)



Voyager took this picture of Ganymede during its 1979 flyby, from about 250,000 kilometers away. Impact craters, icy materials radiating out from them, ridges, grooves, and other surface features are visible down to about a 5-kilometer resolution. (NASA/JPL)

ropa, the satellite just inside Ganymede, goes around Jupiter twice, and Io, the moon inside Europa, goes around four times. Thus, during every orbit the three satellites are aligned, magnifying the gravitational effect. This increased gravitational pull and then relaxation not only causes the orbits to become elliptical but also causes stresses within the satellites themselves. This tidal flexing generates heat that melts ice and causes the surface of Ganymede to be smoother than expected. Many of Ganymede's impact craters have had their depths reduced by the changing surface of Ganymede.

The surface of Ganymede is a mixture of bright areas and dark regions. The dark regions are heavily cratered; the bright areas are less cratered but are often grooved. The bright areas consist of the soil left after the ice has sublimed away. The surface appears much like the surface would if there were lava flow. Water melted by the tidal flexing, then freezing as it reaches the surface, may still flow, much as glaciers flow on Earth. The heavily cratered regions are older and often show fractures and grooves through the craters, where water has broken through the surface at a crater. Ice on Ganymede causes the reflection of radar to be much greater than on most other satellites. Ice also allows radar to penetrate more deeply into Ganymede than if the surface were all silicates. The percentage of ice has been measured at 45-55 percent.

The bulk density of Ganymede is between that of ice and that of carbonaceous silicates, indicating a mixture of the two materials. Ganymede has differentiated; that is, the components have separated, producing a core of dense metals, probably iron or iron with sulfur. A molten iron core is usually the reason for a magnetic field, which Ganymede does have. One model, which agrees with the measured values, postulates a large core of iron with 10 percent sulfur. The moon has a core with a radius of 695 kilometers, a silicate mantle, and a 900-kilometer-thick ice-water shell. Bombardment by meteors causes a change in the albedo of Ganymede in two ways. First, the meteor causes new, darker material to be thrown up onto the surface (the underlying silicates are darker than the ice or residue left by subliming ice). Second, meteor impacts cause the loss of volatile material, leaving an opaque, dark material.

Ganymede has a thin atmosphere, composed of electrically charged gases. One gas seen in the space around Ganymede is hydrogen. Water sublimed from the surface or escaping from a surface fracture condenses at the poles, producing a whitish polar cap down to latitudes of about 40°. Near-infrared spectra show the expected water and hydrated minerals. Unexpected are the indications of carbon dioxide, carbon bonded to hydrogen, carbon triple-bonded to nitrogen, sulfur bonded to hydrogen, and sulfur dioxide. The carbon dioxide appears to be trapped in the surface, perhaps in small bubbles. Jupiter's magnetic field causes ions to be swept along the orbit of the satellite, generating a current producing an auroral spotlight onto the poles of Jupiter.

Ganymede has an intrinsic magnetic field that is opposite to the field of Jupiter. It also displays an induced magnetic field caused by the strong rotating, angled field of Jupiter. The induced field is an indication of a conducting ocean deep under the icy surface. If the ocean has enough minerals dissolved in it to make it strongly conducting, it could generate the intrinsic magnetic field. Jupiter's strong magnetic field causes Ganymede to be bombarded by charged particles. This bombardment is thought to cause the molecular oxygen, O_2, and ozone, O_3, found in the surface of Ganymede.

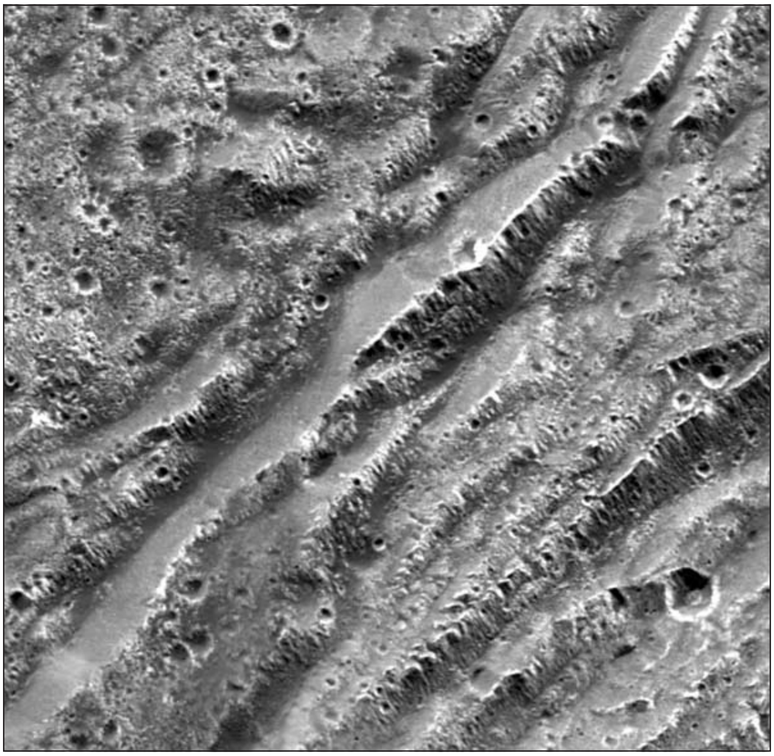
Since Ganymede's orbit is in the same plane as Jupiter, it is thought that they were formed by the same process. Ganymede is out away from the very hot, dense region where the planet formed. Ganymede was formed in a cooler region, where water did not boil away but instead froze to form part of the satellite. Although Ganymede is locked into the same face toward Jupiter all the time, there are indications that this may not have always been true. One clue is that the number of meteor craters should be greater on the leading side of Ganymede, as is the case with Callisto, but this is not true of Ganymede. Another fact pointing to a change in the part of the ice shell facing Jupiter is the catenae that are found on the back side of Ganymede. Catenae are caused by a string of fragments from a comet that was broken up by the intense magnetic field of Jupiter but escaped capture to hit one of the satellites. They should occur only on the Jupiter-facing side of Ganymede.

KNOWLEDGE GAINED

The Jovian system, as Jupiter and its moons are called, has been visited by several space missions. Pioneer 10 (1973), Pioneer 11 (1974), Voyager 1 and Voyager 2 (both in 1979), and New Horizons (2007) flew through the system. They all used gravitational assists to gain speed to travel on toward the outer part of the solar system.

The Pioneer spacecraft provided the first visual images of the surface of Ganymede, as well as a much better estimate of the size and mass of Ganymede. The Voyager spacecraft improved these images to a resolution of a kilometer and provided color by means of six filters. Scientists were able to develop ideas of how the satellite formed and its structure. The Voyager data on craters indicated either that Ganymede's surface has changed, and thus erased the early craters from meteor hits, or that the surface was not firm enough to retain the early craters.

The Galileo mission arrived in orbit about Jupiter in 1996. The visual camera increased



The Galileo spacecraft flew by Ganymede more than two decades after Voyager, capturing this image of dark, heavily cratered terrain and scarps down to a resolution of about 20 meters. (NASA/JPL/Brown University) the resolution of the surface to 20 meters. The model of the moon, showing its core, mantle, and shell of ice, was developed after data from Galileo provided Ganymede's mass, average density, and moment of inertia. The moment of inertia and average density required the data on gravitational fields produced by Galileo, and how the flight path was perturbed as the craft flew by the satellite. Galileo also provided information of the magnetic fields using a magnetometer. The dual purpose of the magnetometer was to determine if Ganymede had a magnetic field of its own and how the satellites interacted with the strong magnetic field of Jupiter. Galileo used its Near Infrared Mapping Spectrometer (NIMS) to make a compositional map of the surface.

While studying Jupiter was not the main focus of New Horizons, astronomers did not let the opportunity escape. Although as of 2008 only 70 percent of the data had been transmitted to Earth, and only part of those data had been analyzed, New Horizons added some interesting new information. The spacecraft's infrared Linear Etalon Imaging Spectral Array (LEISA) and its panchromatic Long-Range Reconnaissance Imager (LORRI) charge-coupled device camera mapped Ganymede's composition. These instruments' resolutions are better than any land-based instrument or Galileo's NIMS. Low-temperature crystalline ice was found as expected, but asymmetric bands of non-ice were found, especially in the darker regions. More ice is found in bright regions and in craters and ejecta from recent meteor hits. This correlates with darker material on the surface, except where meteor strikes have brought ice to the surface. New Horizons could also map parts of Ganymede that Galileo could not see.

Not all information is gathered by spacecraft. The Hubble Space Telescope (HST) has taken pictures of the auroras of Jupiter. Other types of data, such as those gathered by eclipse radiometry, can be used from Hubble or from Earth. Eclipse radiometry is the measurement of thermal radiation just as the satellite is eclipsed by the planet. For Ganymede, these studies suggest that heat is lost rapidly; therefore, the surface material must be porous, due to bombardment from meteors over millions of years.

CONTEXT

The more astronomers learn about large bodies like Ganymede, the more is revealed about how the solar system was formed and about Earth and its Moon. Ganymede may be showing the action of plate tectonics. Learning about the plate tectonics of Ganymede may explain what happened on Earth as the continents tectonically rearranged.

Each space mission has returned valuable information on how to survive in space. Not only are meteorites a danger but gravitational wells, and especially strong magnetic fields, can damage a spacecraft. When humans venture forth, all of those dangers will have to be considered. The number of craters on Ganymede gives scientists some indication of the chance of a meteor hitting the Earth.

C. Alton Hassell

FURTHER READING

Asimov, Isaac, and Richard Hantula. *Jupiter*. Milwaukee, Wis.: Gareth Stevens, 2002. The famous science-fiction author covers the planet and its satellites, the space missions that have studied them, and the comet collisions of 1994. Illustrations, bibliography, index.

Corfield, Richard. *Lives of the Planets*. New York: Basic Books, 2007. The author takes the reader through the different space missions. Divided by planets, the information gathered by each mission is discussed. Index.

Fischer, Daniel. *Mission Jupiter: The Spectacular Journey of the Galileo Spacecraft*. New York: Copernicus Books, 2001. The author takes the reader through each step of the journey of this amazing space probe. Illustrations.

Grundy, W. M., et al. "New Horizons Mapping of Europa and Ganymede." *Science* 318 (2007): 234. This article is one of the first published after the flyby of Jupiter by the Pluto-bound New Horizons spacecraft. Illustrations, bibliography.

Leutwyler, Kristin. *The Moons of Jupiter*. New York: W. W. Norton, 2003. Ganymede is covered in one of the main sections of this book. Illustrations, index.

McFadden, Lucy-Ann Adams, Paul Robest Weissman, and T. V. Johnson, eds. *Encyclopedia of the Solar System*. San Diego: Academic Press, 2007. The editors have collected articles written by many experts in one of the best scholarly surveys of material about the solar system. Illustrations, appendix, index.

Slade, Suzanne. *A Look at Jupiter*. New York: PowerKids Press, 2008. Written for the juvenile audience, this book covers the important information in an easy-to-read style. One big section is devoted to Ganymede and Callisto. Illustrations, bibliography, index.

See also: Callisto; Europa; Io; Jupiter's Magnetic Field and Radiation Belts; Jupiter's Ring System; Jupiter's Satellites; Neptune's Satellites; Planetary Satellites; Saturn's Satellites; Uranus's Satellites.

General Relativity

Category: The Cosmological Context

The general theory of relativity describes the effects of acceleration and gravity on bodies, as well as the structure of space and time. Developed by Albert Einstein in 1915, it supersedes Isaac Newton's laws of mechanics and gravitation. Although Newtonian physics still is sufficiently accurate in many situations in astronomy and physics, general relativity is needed for a deeper understanding of the universe.

OVERVIEW

The general theory of relativity is a model of how gravity works. Gravity is a major force in the universe. It holds stars, the solar system, and galaxies together. It affects the expansion of the universe as a whole. Any model of the universe, then, must have its basis in a theory of gravity. The general theory of relativity has been confirmed in repeated observational and experimental tests to become the basis of modern cosmological theory.

The motivation of Albert Einstein (1879-1955) in working out the principles of relativity was the belief that it is impossible to detect motion relative to any fixed point in space and that, therefore, there is no absolute motion. The general theory is the second of two stages of a larger picture of relativity theory. The first stage is special relativity, which deals with the laws of physics as seen by observers in uniform, or unaccelerated, motion. The general theory goes beyond the special theory to deal with accelerated motion and gravity.

General relativity describes the universe in terms of four dimensions: three dimensions of space, with time as a fourth dimension. Events occur at specific locations in space and at specific moments of time, thus requiring four coordinates (three of space and one of time) to identify each event uniquely. In general relativity, the four-dimensional fabric of the universe is referred to as space-time.

A fundamental concept in the general theory of relativity is the principle of equivalence, which asserts that the effects of acceleration and gravity are indistinguishable. This principle is often demonstrated with the use of a thought (or, from the German, "gedanken") experiment. Imagine that a person is in a chamber that has no windows. Now imagine that the chamber is in deep space far from any other mass, so there is no gravitational effect. As long as the chamber is motionless or moving in a straight line at constant speed, the person floats around freely inside, experiencing weightlessness. If a constant force is applied to the chamber, then it accelerates at a uniform rate and the contents of the chamber (the person) are pushed against the side of the chamber opposite to the direction it is accelerating. This accelerating force can be adjusted to match exactly the downward pull that the person would experience if the chamber were on the surface of the Earth. Without a window to determine the motion relative to the outside world, it would be impossible to tell if the force felt was attributable to gravity or to the acceleration of the chamber. From this type of thought experiment, Einstein concluded that gravity and acceleration are equivalent.

Following from the principle of equivalence, Einstein proposed that the Newtonian concept of gravity as a force operating at a distance between individual masses is entirely unnecessary. Instead, he described an alternative way of looking at gravity by asserting that the presence of mass changes the geometry of space-time by curving or warping it. Gravity then is the effect of curved space-time on mass. In the absence of matter, the shape or curvature of space-time is flat. Near massive objects, however, space-time is strongly warped. The larger the amount of matter at any location, the greater is the curvature of space-time around that location. The curvature is greatest near the massive object, and it becomes progressively less with increasing distance.

The curvature of space-time determines the path along which bodies move. In other words, the curvature of space-time causes moving objects to follow curved paths. The Moon, therefore, orbits the Earth not because of the gravitational force between the Earth and the Moon (as explained in classical, Newtonian physics) but because the mass of the Earth curves the space-time around it.

Another thought experiment is useful to picture the curvature of space. Imagine a pool table whose surface is not rigid but is made of a thin rubber sheet. When a large weight is placed on such a pool table, the normally flat sheet stretches, curving around the weight. The heavier the weight is, the more curved is the surface of the pool table. Attempting to play pool with this weight distorting the surface of the table, one finds that balls passing near the weight are deflected from their straight paths. This is a two-dimensional analogy demonstrating how moving objects are deflected by the curvature of space-time around massive objects. The size, shape, structure, and dynamics of the universe as a whole are determined by the net effect of space-time curvature produced by every mass it contains. Since matter and energy are equivalent (as summarized by Einstein's formula $E = mc^2$), energy produces and is affected by space-time curvature the same as mass. The interplay between matter-energy and space-time can be summarized by these two statements:

(1) Matter and energy tell space-time how to curve.

(2) The curvature of space-time tells matter and energy how to move.

While the basic principles of general relativity are straightforward, some of the implications of the theory defy common sense and ordinary experience. Einstein found that classical Newtonian physics, which applies in most circumstances, does not hold when very strong gravitational fields are involved or when velocities approach the speed of light. According to special relativity, for example, uniform motion at speeds approaching that of light affects measurements of length and time. In an accelerated frame of reference or a gravitational field, similar effects take place.

In 1917, two years after Einstein introduced his general theory of relativity, the German physicist Karl Schwarzschild used it to develop his metric equation describing space-time around a nonrotating spherical mass. One consequence of the Schwarzschild metric is that time slows as a result of the curvature of space-time near masses, and experiments have repeatedly shown this effect actually occurs. Any type of clock positioned near a massive object will run slower than an identical clock farther away. Researchers have found, for example, that clocks in Boulder, Colorado, a mile above sea level, gained about fifteen-billionths of a second per day as compared with clocks near sea level. The difference is attributed to Boulder's slightly greater distance from the Earth's center.

Another consequence that has been confirmed by observation and experiment is that the wavelength, frequency, and photon energy of electromagnetic radiation (light, radio waves, gamma rays, and so on) are affected by the space-time curvature around masses. Electromagnetic radiation has both wave-like and particle-like properties. As a wave, it can be characterized by wavelength and frequency. As particles, called photons, it can be characterized by the energy of the photons. Wavelength and frequency are inversely related, and photon energy is proportional to frequency and inversely proportional to wavelength. For example, radio waves have the longest wavelengths, the lowest frequencies, and the lowest photon energies. Blue light has a shorter wavelength than red light, and thus a higher frequency and higher photon energy. Photons falling or climbing through a gravitational field gain or lose energy, and thus are shifted to shorter or longer wavelengths. For example, if visible light is emitted from the surface of a very massive object, then its wavelength will be lengthened and the color of the light will move toward the red (longer wavelength) end of the visible spectrum as seen by a distant observer. Careful measurements of the wavelength or photon energy of light emitted by and climbing upward through the gravitational field of white dwarf stars and of gamma rays falling through the Earth's gravitational field have confirmed that this shift actually does take place.

APPLICATIONS

General relativity predicts that photons follow curved paths through curved space-time near masses. This was confirmed during a total solar eclipse in 1919. Stars were photographed near the edge of the darkened eclipsed Sun, and the same star field was photographed six months later in the night sky. The positions of the stars in the two photographs were compared, and it was found that their apparent positions near the eclipsed Sun were shifted by an amount consistent with the predictions of general relativity. When seen near the Sun during an eclipse, the positions of stars appear to shift because the path of their light is bent as it passes near the mass of the Sun. This bending of light in a gravitational field is also predicted classically, but the classical effect is only half as large as the relativistic effect. Repetitions of this test during many later solar eclipses all measured deflections that are consistent with the relativistic value, not the classical one.

The curvature of light in gravitational fields is also responsible for the phenomenon of gravitational lensing. If a relatively nearby galaxy lies almost on a line between Earth and a more distant object, light from the more distant object is deflected as it passes the nearer galaxy, and people on Earth will see one or more displaced images of the distant object. The first such gravitationally lensed image was discovered in 1979, and many more have been found since then.

Another successful prediction of general relativity is related to the orbit of the planet Mercury around the Sun. Mercury's orbit is slightly elliptical, and the whole elliptical orbit slowly precesses (or revolves) around the Sun, so that perihelion (the closest approach to the Sun) occurs slightly later on each orbit. Most of the measured advance of Mercury's perihelion is accounted for by the classical Newtonian gravitational influences of the other planets on Mercury, but a small residual is left over. This was the reason for the prediction that there was an unknown planet, Vulcan, inside the orbit of Mercury. According to general relativity, however, the warpage of space-time near the Sun causes Mercury's orbit to precess. The predicted amount of this relativistic perihelion advance for Mercury agrees within 1 percent with measurements of the actual residual perihelion advance.

A prediction arising from the extension of the wavelength-frequency-energy shift of electromagnetic radiation in gravitational fields surprised and disturbed even Einstein. If a mass were sufficiently compressed, then its gravitational field (in classical terms) or its space-time curvature (in general relativistic terms) would be so strong that photons trying to leave it would lose all their energy, implying that not even light could escape it. At first regarded as a mathematical curiosity of general relativity with no "real-world" application, almost fifty years later this prediction became the basis for describing the nature and properties of black holes. The name "black hole" was coined by the American physicist John Wheeler in 1969 for objects from which electromagnetic radiation cannot escape. Although black holes cannot be observed directly (because nothing, not even light, can escape), they can be detected by their effects on nearby matter.

Black holes usually are categorized according to their mass. Stellar-mass black holes have masses several times to several tens of times the mass of the Sun. This was the first type of black hole to be detected. They presumably form as an end state in the "lives" of massive stars. When massive stars (more than about eight times the Sun's mass) have exhausted all their available ways of generating energy by nuclear reactions, they collapse, rebound, and explode as supernovae. Much of their mass is expelled in the explosion. If what remains is more than about two or three times the Sun's mass, no known force can stop its final collapse becoming a stellar-mass black hole.

Super-massive black holes have masses hundreds of thousands to billions of times the mass of the Sun. Evidence for their existence has been found at the centers of many galaxies, including our own Milky Way galaxy. They probably formed in the early stages of galaxy formation, maybe by collapse of the central part of a protogalaxy, by the collision of lots of stars in the dense center of a galaxy, or by the merger of smaller galaxies.

"Mini" black holes have masses less than that of a mountain or small asteroid. They are still just hypothetical and have not yet been detected. However, they have been suggested as possible solutions to many problems in astronomy. If they exist, they may have been formed by some esoteric process in the aftermath of the big bang.

General relativity has provided the framework for modern cosmological models of the structure and evolution of the universe. It predicted the expansion of the universe before it was actually discovered observationally. Most modern cosmological models start with the big bang, which created space and time, matter and energy. As the universe expanded, galaxies, stars, and planets formed in it. According to general relativity, the average density of matter and energy determine the geometry and future of the universe. The universe appears to have flat geometry, but only a small fraction of the mass needed for this is observed. Most of the mass in the universe seems to be some exotic form of dark matter that has not yet been directly observed but that is beginning to be detected by its gravitational effects. Gravity alone predicts that the expansion of the universe should slow down, yet the expansion appears to be accelerating. The unknown cause of the acceleration has been dubbed dark energy, and there seems to be even more dark energy than dark mass.

CONTEXT

Relativity theory emerged as a logical step forward in the understanding of space and time. It arose out of the inability of Newton's laws of gravity and motion to describe the observed universe in certain circumstances, such as at very high velocities or in the presence of very strong gravitational fields. The general theory of relativity arose out of Einstein's earlier special theory of relativity. The concept of relativity is extremely important for the development of theories about the universe because it is based on two fundamental postulates: All motion is relative, and there is no preferred frame of reference in which space and time are defined absolutely; it is fundamentally impossible to detect motion relative to absolute space, and there is no basic universal time.

Einstein's goal was to develop a unified field theory. He spent the last years of his life searching for the link between gravitation, electromagnetism, and nuclear forces. This remains a major goal of physics, but it has been difficult to find a way to merge the general theory of relativity with quantum mechanics, the theory that describes how subatomic particles interact.

While it is impossible to prove that the general theory of relativity is correct, the observations and experiments that have been used to test it so far have failed to disprove it. It remains the most accurate theory by which to measure and predict the effects of gravity and has become the foundation of modern cosmology and physics. General relativity has altered the way we view space and time. As such it is an important tool in our continuing efforts to describe and predict events in the world and to understand more about the nature of the universe.

Divonna Ogier

FURTHER READING

Barbour, Julian B., and Herbert Pfister, eds. *Mach's Principle: From Newton's Bucket to Quantum Gravity*. Einstein Studies 6. Boston: Birkhäuser, 1995. Covers all aspects of relativity. Looks at historical, philosophical, and theoretical components in addition to experimental approaches. Includes a discussion of Mach's principle at a conference in Germany.

Calder, Nigel. *Einstein's Universe*. New York: Viking Press, 1979. A well-written book directed toward a general audience. Details the principles of special and general relativity in reference to Einstein's life and a general philosophy of the universe. Useful analogies and popular language make this difficult subject more easily understandable.

Chaisson, Eric. *Relatively Speaking*. New York: W. W. Norton, 1988. A highly readable book giving an overview of Einstein's work on special and general relativity and how it relates to modern astronomical and cosmological questions. Gives a nonmathematical analysis of cosmology and singularity theory.

Foster, J., and J. D. Nightingale. *A Short Course in General Relativity*. 2d ed. New York: Springer, 1995. A detailed text geared toward advanced undergraduates and first-year graduate students. Early chapters cover the needed math skills to understand the subject fully, including tensor calculus and differential geometry. Gives solutions to problems and exercises.

Glendenning, Norman K. *Compact Stars: Nuclear Physics, Particle Physics, and General Relativity*. New York: Springer, 1997. Discusses general relativity, nuclear and particle physics, stellar evolution, black holes, neutron stars, and related topics. Also includes several problem sets. For undergraduates and graduate students.

Hawking, Stephen. *A Brief History of Time: From the Big Bang to Black Holes*. New York: Bantam, 1988. The most celebrated physicist of his generation addresses many of the most important concepts in theoretical physics for a general audience. Discusses deep issues at the frontier of physics such as the nature of gravity, the nature of time, and the search for grand unification.

_______. *An Even Briefer History of Time*. New York: Bantam, 2008. An update of Hawking's celebrated first book, *A Brief History of Time*. Discusses string theory, dark matter, dark energy, and the ultimate fate of the universe.

Parker, Barry. *Einstein's Dream*. New York: Plenum Press, 1986. A perspective on relativity theory related to Einstein's ultimate goal of a unified theory of the universe. Gives a good historical analysis of relativity and unification theory. Devotes chapters to the origin of the universe, the ultimate fate of the universe, black holes, and quantum theory. Nonmathematical, but recommended for the reader with a background in physics as well as nonspecialists.

Sartori, Leo. *Understanding Relativity: A Simplified Approach to Einstein's Theories*. Berkeley: University of California Press, 1996. Addresses general relativity and cosmology with an emphasis on Einstein's theory of relativity. Designed for the general reader with basic algebra knowledge, this book provides historical and philosophical background for modern physics courses.

Schwartz, Joseph, and Michael McGuinness. *Einstein for Beginners*. New York: Pantheon Books, 1979. A delightful introductory book for those intimidated by the usual presentations of scientific theory. Written in simple language with comic-book-style drawings, it gives a basic overview of Einstein's life, the political and social environment from which he emerged, and the basic principles of his work.

See also: Big Bang; Cosmic Rays; Cosmology; Eclipses; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; Gravity Measurement; Interstellar Clouds and the Interstellar Medium; Milky Way; Optical Astronomy; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Universe: Evolution; Universe: Expansion; Universe: Structure; White and Black Dwarfs.

Gravity Measurement

Category: Scientific Methods

Gravity, the most dominant universal force, attractive in nature, affects all forms of matter and even energy in spite of its extreme weakness. Traditional Newtonian gravitational theory, adequate for navigation and general astronomical purposes, requires modification when great precision in measurement is necessary.

OVERVIEW

Gravity measurement involves three separate aspects, some theoretical and others experimental or practical. First there is the basic investigation of the nature of gravity as either a force or a consequence of the curvature of space in extreme circumstances, specifically those requiring general relativity rather than classical physics. Second there is the determination of the gravitation field around planetary bodies in order to control satellites in orbit; this is the field of geodesy. Third there is the effort to detect gravitational radiation, with its implications for the relatively new field of gravitational astronomy and also high-energy astrophysics in general.

To the average person, gravity is the natural tendency of objects to move downward toward the Earth. Gravity has been traditionally described as a field with every particle of matter being a source of a gravitational field. The intensity of this field is affected by the distance from and position on the Earth's surface and the local mass distribution in relation to the total mass of the Earth.

The gravitational field produces an attractive force between bodies, which is directly proportional to the product of their masses and inversely proportional to the square of the distance between their centers. This statement summarizes the law of universal gravitation, as formulated by Sir Isaac Newton in 1687. This law assumes that the masses are distributed symmetrically about a sphere of constant radius and uniform density. Actual gravity surveys, however, demonstrate that no mathematical formula has been found that describes exactly the gravitational field of the Earth. The Earth's field is complicated by irregularities in the topography and mass distribution, combined with a pronounced flattening of the Earth at its poles, caused by rotation.

The force of gravity varies with position on the Earth's surface. The acceleration of free-falling bodies caused by the force of gravity is determined experimentally as greatest at the poles and smallest at the equator. The value for the acceleration of gravity, g, for example, is only 9.782 meters per second per second (m/s^2) in the Canal Zone of Panama but is 9.825 m/s^2 in Greenland, which is closer to the North Pole. This value of g near the equator is lessened by a centripetal factor, which is the square of the velocity of a point on the Earth's surface divided by the radius of the Earth (V^2/R). Since points closer to the equator move with a greater velocity, the value of g will be smaller at the equator.

Gravitational acceleration diminishes with altitude; an object at the Earth's surface and near the equator that would have a value of g equal to 9.83 m/s^2 would have that value drop to 8.70 m/s^2 at an altitude of 400 kilometers. This decrease is observed because the effect of the gravitational force on acceleration follows an inverse square law; that is, the farther the object is from the center of Earth, the less the acceleration. At twice the distance, for example, the force and resulting acceleration would be only one-fourth of the original amount.

The equation for the law of universal gravitation has a constant G, the gravitational constant that was not known at the time Newton formulated the law. The constant G is assumed equal for all conditions and locations on the Earth and in the universe. The weight of an object W is equivalent to its mass m multiplied by the acceleration g ($W = mg$). If the weight is equated to the pull of gravity from the law of universal gravitation, then the gravitational constant G may be calculated directly. The value of G is found by squaring the radius of the Earth, R, then multiplying the result by g and dividing by the mass of the Earth, M ($G = R^2 g/M$). This result is a constant for a given location, which implies that g should be the same also for any location. One sees that the surface value of g depends on the mass of a planet or star or other body as well as how the object's mass is distributed within the object. For example, the surface value of g on the Moon is one-sixth that of Earth, and the surface value of g on Mars is 38 percent that of Earth.

The technical problems of the measurement of G were solved by Henry Cavendish in 1798. Cavendish devised a sensitive torsion balance composed of a light rod supported at its center by a thin wire approximately 1 meter long, with lead balls about 5 centimeters in diameter placed at the ends of the rod. If a force is applied to each lead ball in opposing directions and at right angles to both the wire and rod, the wire is subjected to a rotation that may be measured as an angular displacement. Cavendish initially applied small forces, measuring the amount of twisting that resulted. Carefully shielding the experimental equipment from air currents, Cavendish placed two large lead balls about 20 centimeters in diameter nearly in contact with the small lead balls but on opposing sides. Gravitational force between both sets of balls caused a twist in the wire, and from the angle displaced by the wire, Cavendish was able to measure the forces between the large and small balls. The force turned out, as expected, to be very small, only one two-millionth of a newton. The value of G could then be calculated directly, since Cavendish now knew the force involved as well as the masses of the lead balls and their distance of separation. The results of this experiment, as well as later determinations, have established that G has the same value whatever the composition of the masses or the location; the constant is truly universal in nature.

The gravity pendulum has been used to measure the differences in gravitational force on Earth. Modern gravity pendulums are governed by the principle relating to the period of oscillation discovered by the Dutch scientist Christiaan Huygens. The period of the pendulum, as he noted, varies directly with the square root of the length and inversely with the square root of the local value for the acceleration of gravity g. Gravity pendulums are built nearly friction-free, supported on knife-edge jewel bearings, and swing in chambers from which air has been evacuated.

Christiaan Huygens: Improving on Galileo

Christiaan Huygens's first plunge into scientific research took place in 1655, when he and his brother began to build improved telescopes, grinding their own lenses. With these instruments, Huygens found Saturn's moon Titan and discovered that Mars has a varied surface. He gradually discerned a ring around Saturn that nowhere touched the planet, thus improving on Galileo's more primitive observation. In order to protect the priority of this discovery while continuing his viewing, he announced by the publication of a coded message that he had found the ring. At about this time Huygens published his work on hyperbolas, ellipses, and circles, and in 1657 he published the world's first formal treatise on probability.

In addition to building the pendulum clock—his greatest original invention—Huygens enjoyed membership in France's Royal Academy of Sciences, residing in Paris from 1666 until 1681. Still, when Huygens left Paris in 1681 for a third trip to the Netherlands, he never returned. His patron, Louis XIV's chief minister Jean-Baptiste Colbert, died in 1683, and anti-Protestant sentiment was growing in France, making Huygens's position difficult, as he was nominally a Calvinist.

Huygens's philosophy of science was intermediate between those of the two giants of his day: René Descartes in France and Sir Isaac Newton in England. Descartes attempted to explain all phenomena by use of deductive logic alone. Newton, on the contrary, relied on observations and experiments as the bases for his laws. Huygens grew up a Cartesian but broke with his mentor over the latter's extreme devotion to the mathematical, or deductive, approach to science; his basic approach to the universe was mechanistic. He did prefer, however, Descartes's supposedly more tangible "vortices" of "subtle matter" to Newton's "gravity" in explaining the movements of heavenly bodies. In the matter of relativity, however, Huygens was in advance of Newton and anticipated Einstein. For Huygens, all motion in the universe was relative. Huygens also bested Newton in his understanding of light: Newton held to the corpuscular (particle) theory of light, whereas Huygens propounded a wave theory of light; modern quantum theory combines the two, but in Huygens's day the corpuscular theory was dominant.


(Library of Congress)

Huygens remained in communication with Newton, although his relations with London's Royal Society dwindled after 1678. He visited England again in 1689, conversing with Newton and addressing the Royal Society on his non-Newtonian theory of gravity. Huygens's last years were spent in The Hague, where he died in 1695.

The period of oscillation is timed with precise chronometers enabling determinations to within a few parts per million. Unfortunately, gravity pendulums are unwieldy and difficult to transport; consequently, now most gravity measurements are made with portable instruments called gravimeters.

Gravimeters make use of the principle of a spring balance—that the distortion or strain is directly proportional to the applied stress or force, provided that the measurements are made within the elastic limits of the material. A small quartz fiber is distorted in the local gravitational field at an observed station with results compared to a measured pendulum station. The readings are generally so precise that distortions to one part in 10 million can be recorded. Through development of such instruments, variations in gravity over large areas can now be measured. Some instruments have been adapted to operate from aircraft in flight for aerial surveys of the Earth and for use on surface vessels at sea in regions not suitable for gravity pendulums because of wave disturbances and motions.

A torsion balance employs an arrangement similar to a Cavendish balance, but as opposed to measuring the deflection produced by large masses, the period of oscillation is measured when the large masses are placed perpendicular to the equilibrium positions of the small masses and next when the large masses are rotated another 90°. When the large masses are located at the first position, the period is less as a result of the additional restoring force. In the second position, the period is increased by the large masses, pulling the small ones away from the equilibrium position. The difference in the periods from both of these measurements gives the value of G. The advantage of this method is that the period may be measured more precisely than a corresponding deflection; the precision in error is estimated as 0.5 percent with this technique.

Although measurement of the gravitational constant G has improved, the value is not known with nearly the precision of other physical constants. Laboratory measurement of G is difficult because of the extremely small forces between the masses. Planet-sized objects are much larger, but the problem is not resolved, because the product of G and the mass of the attracting planet both appear in the equation. Planetary observations alone cannot determine the individual values of G or mass.

A torsion balance may also be used for ascertaining the equivalence between inertial mass and gravitational mass, known also as the principle of equivalence. For these experiments, one body, called the inertial mass, is defined with respect to a standard mass of 1 kilogram. The gravitational mass of the body is defined in terms of this acceleration and the distance between the objects. Experiments have tested a variety of materials, obtaining a ratio between the masses. Results indicate that various types of energy contribute to the inertial mass of a system to the same degree that they would contribute to the gravitational mass.

A gravity gradiometer is an instrument designed to measure local tidal fields. The instrument is portable and designed for use on an airplane or satellite and permits precise mapping of anomalies in the Earth's gravitational field. The instrument, in the shape of a Greek cross, has four masses at the ends of each arm, which are held together at the center by a torsional spring. When pressed together, the arms oscillate with a frequency of 32 hertz (Hz, or oscillations per second). If placed in a tidal field at right angles, the cross will be deformed. Rotated with an angular frequency of 16 Hz in the reference frame of the cross, a tidal driving force would appear at 32 Hz. Since the oscillation frequency matches the natural vibration frequency of the arms, a resonance condition is established, producing large amplitudes. Very small tidal fields have been detected with this instrument.

Detectors of gravitational radiation, or waves, were first built in 1966 at the University of Maryland by Joseph Weber. This type of detector consists of a large aluminum cylinder placed inside a vacuum tank and suspended on a wire. The cylinder is supported on rubber blocks as an insulation to external mechanical vibrations. Any oscillations of the cylinder in the fundamental or longitudinal mode are detected by piezoelectric strain transducers bonded to the outside middle section. The ability of the device to detect radiation at resonance is termed its cross section, which turned out to be quite small. Limiting the ability to detect this type of radiation is the thermal motion of the individual molecules and oscillations of the cylinder that interfere with observations. Experimental results obtained by Weber and his group have not been duplicated elsewhere, casting doubt on the technique's reliability.

More sensitive detectors were constructed and involved strains in solid bodies employing resonance; they were more sensitive to radiation of a given frequency and rejected all other frequencies. The biggest advance in gravitational wave detectors came with approval of large-scale facilities for detection, one on Earth (called LIGO) and one based in space (called LISA).

The Laser Interferometer Gravitational-Wave Observatory (LIGO) consists of two facilities working in coordination, one at Livingston, Louisiana, and another on the Hanford Nuclear Reservation in Washington State. The two sites are separated spatially by 3,002 kilometers and temporally by approximately 10 milliseconds, since gravitational waves would travel at the speed of light. Separating the facilities permits an interferometry approach for determination of a source of gravitational radiation. Gravity wave detection with LIGO involves looking for oscillations in lasers traveling down vacuum lines to mirrors. Essentially a higher-technology version of the classic Michelson interferometer, LIGO was designed to be able to detect a spatial oscillation of as little as 10^{-18} meter when a gravitation wave forces an oscillation as it passes through Earth. Note that this tiny spatial extent is approximately one-thousandth the size of a proton or neutron. If detected, a gravitational wave would produce a signal in LIGO that rises rapidly in oscillation amplitude, only to decay exponentially in time. An early test of LIGO was an observation of the gamma-ray burst GRB070201 in 2007. Although near the direction toward the neighboring Andromeda galaxy, the burst was determined not to have come from within that galaxy. Its closeness is such that LIGO should have picked up gravitational radiation if the object emitting the intense gamma-ray burst had actually been the Andromeda galaxy, but it failed to pick up any signal.

Another gravitational radiation detection project under development is the Laser Interferometer Space Antenna (LISA). LISA will consist of three separate Michelson interferometers, each with two long arms arranged in the shape of the letter L positioned precisely in space about 50 million kilometers from Earth. The system, to be built by the National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA), is planned for launch no earlier than 2018 for a minimum two-year mission. LISA's design includes a capability for determining spatial oscillations as small as 1 part in 10^{20}.

APPLICATIONS

Measurement of the gravitational field over Earth demonstrates that the field is not uniform. The Earth departs from being a perfect spheroid because of rotational effects and topographic variations. Regions that are topographically higher than a datum surface are located farther from the Earth's center and experience a smaller gravitational force. Other regions located below the surface, although closer to the Earth's center, may experience compensating effects from mass concentrations, thereby increasing the strength of the field over its expected value.

Gravity data indicate that the field is increased near mountain ranges because of the greater concentration of rock. Closer observations show that mountain masses do not deflect the field as much as expected if the mountain were a load resting on top of a uniform crust. If the mountain were merely a load on the rigid crust, the force of gravity (corrected for the effect of additional altitude) should be larger on top of the mountain than on the surrounding plains as a result of the increased gravitational pull of the mountain mass beneath the crust. Such observations led geophysicists to conclude that a rigid crust is not responsible for supporting the load of the mountains but is instead buoyed up by floating on a denser, deformable interior. The interior of the Earth must yield and be subject to lateral flow to compensate for loads on equal-size regions. Areas of depression in the crust, such as oceanic trenches, show lower values of the gravitational field, as there is less mass near the surface.

Newtonian mechanics, traditionally used to describe the behavior of bodies at the surface of the Earth, tend to break down outside the range of normal observable motion. The theory fails when gravitational fields become very intense near collapsed objects such as neutron stars or black holes. In 1915, Albert Einstein completely changed the understanding of gravitation with his general theory of relativity. According to theory, gravity is not a force in the usual sense but is the result of the curvature of space-time. Bodies then follow the easiest course through space-time, which is manifested in the shape of their orbits. Einstein theorized that gravity may be explained by geometry. Mercury's orbit could not be explained adequately by ordinary mechanics but only by the warping of space near the Sun. Time warps at the Earth's surface may be detected by using very precise clocks.

Gravitational time dilation—the slowing down of clocks in a gravitational field—may be used as a direct test of the curvature of space-time. For these experiments, cesium-beam atomic clocks are used. Small frequency shifts are measured in the clocks placed in a gravitational potential and are calibrated against clocks at rest in a stationary gravitational field. The pulses of the clocks are monitored to rule out the possibility of a frequency loss during the light-beam propagation. The clock tick rate is found to depend upon the strength of the gravitational field; therefore, space-time geometry is dependent upon the gravitational field. The possibilities that strong tidal forces would have an effect on the clock may be ruled out, because it is known that the atomic forces are stronger and resist tidal distortions.

Gravitational redshifts have been measured on light emitted from the atoms on stellar surfaces. It has been difficult to obtain reliable results for these measurements because of strong convection currents in the stellar photosphere masking the spectral lines, which are Doppler-shifted by gaseous motion. Measurements made above the photosphere of the Sun have given more definitive results. Gravitational redshifts are very prominent in light emitted by white dwarf stars because these stars have about 1 solar mass contained in a much smaller radius than the Sun and consequently have very intense gravitational fields.

LIGO was used to study the pulsar in the center of the Crab nebula. That pulsar's spin rate (30 revolutions per second) is slowing down considerably; indeed it is one of the most quickly slowing pulsars known. Pulsars slow down in their periodic spinning motion due to one of three mechanisms: (1) magnetic dipole radiation emission, (2) asymmetric emission of energetic particles, and (3) possible emission of gravitational radiation. General relativity indicates that if a rapidly rotating neutron star, a pulsar, were not perfectly smooth but distorted from a spherical shape, it should create gravitational waves. Since LIGO picked up no signals within detection capabilities during Crab pulsar observations, the conclusion was that to within well under a meter, the pulsar must be nearly perfectly spherical.

CONTEXT

Electromagnetic waves were a prediction of James Clerk Maxwell's field equations for electric and magnetic interactions. Heinrich Hertz verified that prediction's correctness with the first detection of radio waves; prior to that point, the only portions of the electromagnetic spectrum known, aside from visible light, were parts of the ultraviolet and infrared. In due course, astrophysics would be able to make use of observations across the electromagnetic spectrum, from high-energy gamma rays to the lowest-energy radio waves.

Gravitational waves are predicted by Einstein's field equations of general relativity. However, whereas production and detection of electromagnetic radiation came in rapid succession and opened up a vast expansion of astronomical investigations as a by-product, prediction and initial direct detection of gravitational radiation remain separated greatly in time. Despite null results, gravitation physicists remain steadfast in their belief that gravitational radiation should exist and that, in time, new technology such as LIGO and LISA will make that first definitive direct detection of such radiation. Thus far, the existence of gravity waves has been only indirectly inferred by examining the binary pulsar system PSR 1913+16, for example, and other results have been used to infer aspects of other high-energy exotic objects in the universe. When and if the technology exists to observe and examine gravitational radiation routinely, a brand-new branch of astronomy will provide a host of new insights into the nature of the universe and one of its most pervasive forces, gravity, which binds it together into complex structures and in part determines its evolution.

All of nature's events and activities can be explained in terms of four fundamental forces. In the historical context, gravity was the first of these four forces to be investigated scientifically. Although scientists have had an awareness of gravity and the direction in which it acts, the role of gravity as a force was not appreciated fully until Newton's law of universal gravitation was published. The importance of gravity is its universal nature. Everything in the cosmos is affected by it and every particle of matter is a source of gravity. The force of gravity as observed is always attractive, tending to pull matter together.

One of the surprising facts concerning gravity as the dominant universal force is its extreme weakness. Gravity is so weak that physicists generally ignore its effects completely when dealing with masses on the level of the subatomic particle. Gravity's strength on the atomic scale is vastly overwhelmed by the electromagnetic and nuclear forces at that distance.

The law of universal gravitation, which was adequate for more than two hundred years, was not effective in the twentieth century in explaining discrepancies in observations near very massive objects. In this respect, the law of universal gravitation conflicted with the relativity theory. In Newton's theory, gravitational force between two bodies should be transmitted instantaneously across space, but Einstein's theory rejects physical effects that travel faster than the speed of light. Gravitational fields around objects as massive as the Sun appear to distort space and time to a degree that is detectable. Observing stars near the Sun during solar eclipses indicates that they are not observed in their true positions; that is, the light from these stars has been bent or deflected noticeably toward the Sun by its gravitational field. Black holes appear to distort severely the surrounding space and time. They are the final state of very massive stars that have collapsed into a singularity with a gravity so overpowering that not even light can escape.

Some physicists have postulated the existence of a fifth force in nature that may diminish the effectiveness of the gravitational force out to a limited range. Experiments performed in mines, for example, seem to show that the measured gravitational force does not agree with predicted values from theory. These observations as well as others, however, have not established conclusively the existence of a previously unknown repulsive force.

Despite all of its success, general relativity theory remains at odds with quantum theory (which describes subatomic particles on a statistical basis) and is at odds with superstring theory (which treats subatomic particles as very tiny vibrating loops). Physicists continue to search for a more comprehensive theory of quantum gravity that perhaps will be more useful in mapping out the very early history of the universe very near the moment of creation. The nature of just how gravitation is transmitted at a distance has not been resolved as to whether it has a wave or particle nature, or both. If gravity has a particle nature, then this particle must be extremely tiny because gravity as a force is very weak. Unification of gravity with the other three fundamental forces remains the biggest central problem in theoretical physics. After more than three decades, string theory has failed to achieve that unification. However, a large percentage of the physics community remains committed to string theory over other avenues of approach to the problem.

Michael L. Broyles

FURTHER READING

Beutler, G., M. R. Drinkwater, R. Rummel, and Rudolf von Steiger. *Earth Gravity Field from Space: From Sensors to Earth Sciences*. New York: Springer, 2003. Tells the story of NASA's GRACE satellite mission, in which gravity field measurements improved knowledge of Earth's interior and ocean dynamics. Technical.

Fowles, Grant R., and George L. Cassiday. *Analytic Mechanics*. 7th ed. New York: Brooks/Cole, 2004. A college textbook for a second course in Newtonian mechanics. Particularly strong on theory and applications of orbital motion. Requires knowledge of advanced calculus.

Gamow, George. *Gravity*. New York: Dover, 2003. A famed physicist in his own right, Gamow provides insight into the thinking of Galileo, Newton, and Einstein that led to contemporary understanding of the nature of the space-time continuum. Explains gravity as a consequence of the curvature of space. Accessible to the general reader.

Greene, Brian. *The Elegant Universe: Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory*. New York: W. W. Norton, 2003. A rather remarkable attempt to explain the most esoteric aspects of string theory and quantum field theory for the interested general reader. Some chapters require intense attention and multiple readings. Later chapters go into far more rigorous mathematical detail than the general reader may be able to follow.

Hawking, Stephen. *A Brief History of Time: From the Big Bang to Black Holes*. New York: Bantam, 1988. Hawking, perhaps the most celebrated physicist of his generation, provides a highly readable text that addresses many of the most important concepts in theoretical physics. Discusses deep issues at the frontier of physics, such as the nature of gravity, the nature of time, and the search for grand unification.

_______. *An Even Briefer History of Time*. New York: Bantam, 2008. An update to Hawking's celebrated *A Brief History of Time*, in which the physicist discusses string theory, dark matter, dark energy, and the ultimate fate of the universe.

Hawking, Stephen W., and William Israel. *Three Hundred Years of Gravitation*. New York: Cambridge University Press, 1987. A comprehensive treatise spanning the development of gravitation from Newton to concepts of quantum gravity and time asymmetry. Additional chapters discuss gravitational radiation, gravitational interaction of cosmic strings, inflationary cosmology, quantum cosmology, and superstring unification. An important reference in gravitational physics.

Hofmann-Wellenhof, Bernhard. *Physical Geodesy*. New York: Springer, 2006. An update to a text originally published four decades earlier. Offers thorough explanations of satellite methods for determining Earth's gravitational field. Highly mathematical.

Kaula, William M. *Theory of Satellite Geodesy: Applications of Satellites to Geodesy*. New York: Dover, 2000. A reprint of a classic book that was used by a generation of researchers involved in gravity field modeling. Technical.

Parker, Sybil P., ed. *McGraw-Hill Encyclopedia of Physics*. New York: McGraw-Hill, 1983. An excellent reference for the nontechnical as well as for the technical reader. Topics related to gravitation include Newton's law of universal gravitation, gravitational constant, mass and weight, gravity, gravitational potential energy, application and accuracy of Newtonian gravitation, relativistic theories, supergravity, and gravitational waves.

Thornton, Stephen T., and Andrew Rex. *Modern Physics for Students and Engineers*. 3d ed. New York: Brooks/Cole, 2005. A comprehensive presentation of the development of relativity, quantum mechanics, and nuclear and particle theory and experimentation. For undergraduate science majors and serious scientific researchers.

See also: Archaeoastronomy; Coordinate Systems; Earth System Science; General Relativity; Hertzsprung-Russell Diagram; Infrared Astronomy; Neutrino Astronomy; Optical Astronomy; Planetary Orbits; Planetary Orbits: Couplings and Resonances; Radio Astronomy; Solar Geodesy; Telescopes: Ground-Based; Telescopes: Space-Based; Ultraviolet Astronomy; X-Ray and Gamma-Ray Astronomy.

Greenhouse Effect

Category: Earth

"Greenhouse" gases absorb or trap infrared radiation emitted by a planet's surface. The absorbed or trapped energy is then released or reemitted by the greenhouse gases, resulting in additional heating of the planet's surface. Greenhouse gases include water vapor, carbon dioxide, methane, nitrous oxide, ozone, and a class of human-made molecules called chlorofluorocarbons. There is international concern that increasing atmospheric levels of these gases could lead to a global warming.

OVERVIEW

The Earth and the other planets in the solar system receive almost all their energy in the form of electromagnetic radiation from the Sun. While the Sun emits radiation over the entire electromagnetic spectrum, from shortwave X rays to longwave radio waves, the bulk of the solar radiation is in the ultraviolet, visible, and infrared parts of the electromagnetic spectrum, from about 0.15 to about 4 microns, or about 150 to about 4,000 nanometers. The peak of the Sun's radiated energy is in the visible part of the spectrum, at about 0.5 micron, or 500 nanometers, as that is the wavelength of maximum emission for an object at a temperature of about 6,000 kelvins, which is the temperature of the Sun's "visible surface," the photosphere. The intensity falls off toward shorter (bluer) and longer (redder) wavelengths, and the human eye perceives this distribution of intensity versus wavelength as a yellowish-white color. Hence, the Sun appears as a yellowish-white object in the sky.

The amount of solar radiation intercepted by a planet and potentially available for heating it depends on the Sun's luminosity, the planet's distance from the Sun, and the planet's cross-sectional area. For example, at Earth's distance from the Sun, about 150 million kilometers, the flux of solar radiation, also called the solar constant, is 1,368 joules per second per square meter multiplied by Earth's cross-sectional area of 1.28×10^{14} square meters; thus, the solar energy intercepted by Earth is about 1.75×10^{17} joules per second. Not all this intercepted solar radiation, however, is available for heating the planet's surface; some of the solar radiation is reflected back to space by the planet's surface and by clouds if the planet has an atmosphere.

The fraction of solar radiation reflected back to space is called the albedo of the planet. The solar radiation not reflected is absorbed by the planet and heats its surface. (To continue our example for Earth, its albedo is about 33 percent. Therefore, about 67 percent of the intercepted solar radiation is not reflected but instead is absorbed, heating Earth's surface.) The planet's surface in turn emits its own radiation at the same rate as it absorbs solar radiation. If the planet emitted radiation at a rate less than the rate it absorbed solar radiation, the planet would heat up and emit more radiation. However, if the planet emitted radiation at a rate greater than the rate at which it absorbed solar radiation, the planet would cool down and emit less radiation. The planet's rate of emitting radiation depends on its temperature. Equating the planet's rate of emitting radiation to the rate it absorbs solar radiation yields the planet's "effective temperature," which is the temperature of an ideal thermal radiator (or blackbody) that emits radiation at the same rate as the planet.

Earth's effective temperature is calculated to be 253 kelvins (–20° Celsius). At this temperature, Earth emits radiation in the infrared part of the electromagnetic spectrum, most of it between about 4 and 80 microns. Most of the incoming solar radiation with wavelengths between about 0.3 and 2 microns travels through the atmosphere without significant attenuation or absorption by atmospheric gases, and the infrared radiation flowing outward from Earth is absorbed by several atmospheric gases (mainly water vapor and carbon dioxide), called greenhouse gases (GHGs). After a fraction of a second, the GHGs reemit the infrared radiation randomly in all directions. About half of the reemitted infrared radiation is directed outward and about half is directed downward, back toward the surface. The downward component of the reemitted infrared radiation is absorbed by the surface, with an additional heating effect. Hence, Earth's surface is heated not only by incident solar radiation (wavelengths of 0.3 to 2 microns) but also by Earth-emitted infrared radiation (wavelengths of 4 to 80 microns) that was absorbed and then reemitted by atmospheric GHGs. The additional heating of the Earth's surface by the reemitted infrared radiation heats the surface about an additional 35 kelvins, raising the effective temperature of 253 kelvins to the actual average temperature of about 288 kelvins, or 15° Celsius. The temperature enhancement of 35 kelvins is termed the greenhouse effect. It is this temperature enhancement that makes the Earth habitable for life; without it, our average temperature would be well below the freezing point of water.

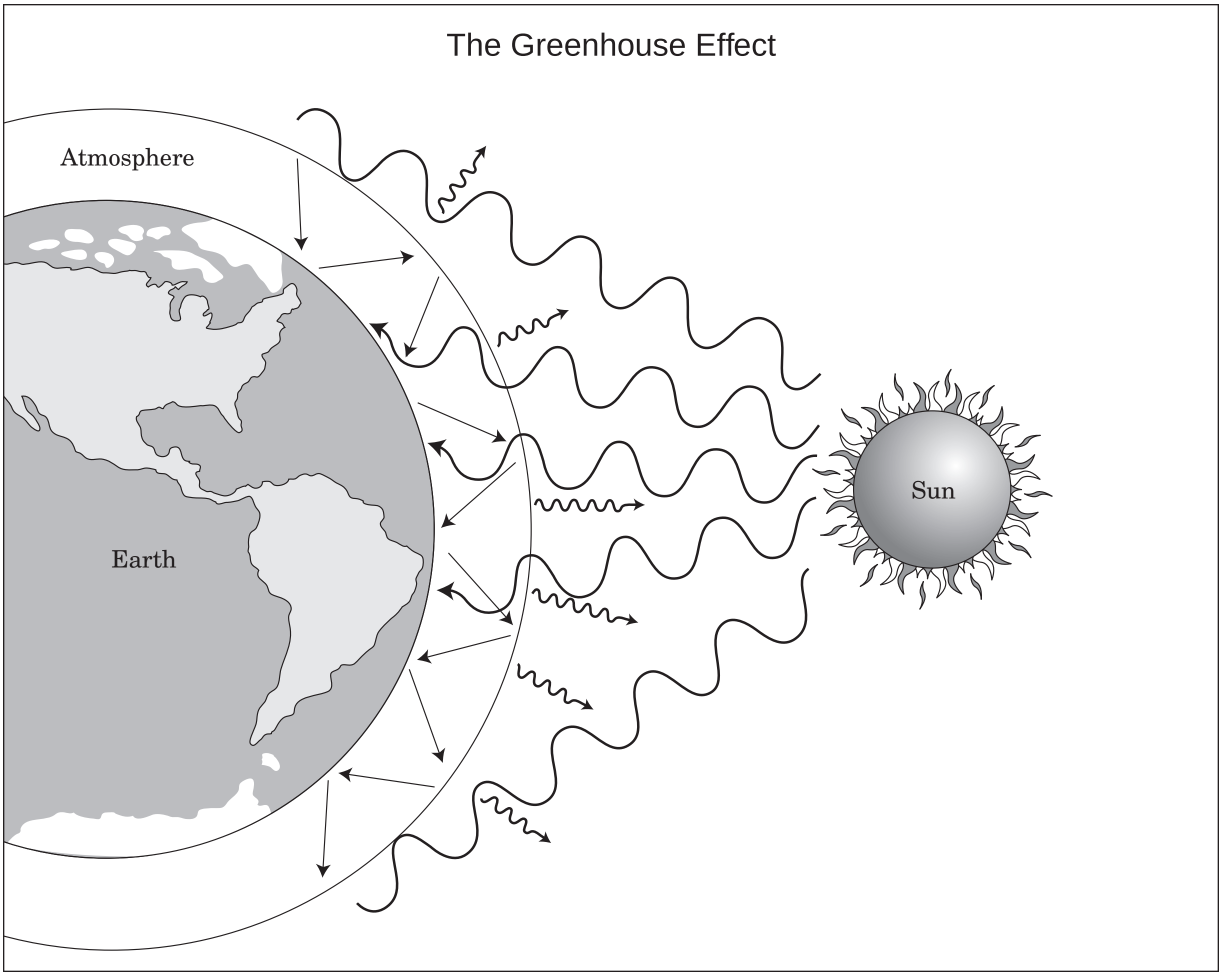
Venus is closer to the Sun than the Earth is, about 0.7 times our distance, but its albedo is about 0.65, meaning it reflects much more sunlight than Earth does; as a result, Venus appears very bright in the night sky. Consequently, the effective temperature of Venus is nearly the same as Earth's; it receives more solar radiation, but it also reflects more and thus absorbs less of it. However, while GHGs are only very minor constituents of the Earth's atmosphere, carbon dioxide is the major constituent of the atmosphere of Venus. About 96 percent of the molecules in Venus's atmosphere are carbon dioxide, with nitrogen accounting for the remaining 4 percent.

The surface pressure of the atmosphere of Venus is about 90 times the pressure of our atmosphere at sea level. As a result of the high percentage of carbon dioxide and the high surface pressure of the atmosphere, there is a very efficient and significant greenhouse effect on Venus. The greenhouse effect increases the temperature of Venus by about 480 to 500 kelvins, from its effective temperature of about 250 kelvins (close to the effective temperature of Earth) to the measured surface temperature of Venus of about 730 to 750 kelvins (about 855° to 890° Fahrenheit)—hot enough to melt lead.

There is international concern that the buildup of GHGs in Earth's atmosphere could lead to an enhanced greenhouse effect and global warming, resulting in record-high temperatures in many regions and severe droughts in others. Such global warming would lead to thermal expansion of the oceans, since water expands in volume when heated. More significantly, global warming could result in substantial melting of the ice and snow in Earth's polar regions, adding an even greater volume of water to the world's oceans. A rise in sea level would result in the flooding of many low-lying land areas.

The most important atmospheric GHGs are water vapor, carbon dioxide, methane, nitrous oxide, tropospheric ozone, and a human-made family of gases termed chlorofluorocarbons (CFCs), of which CFC-11 and CFC-12 are the most abundant species. All the GHGs are very minor constituents of Earth's atmosphere, which is composed primarily of nitrogen (78.08 percent by volume), oxygen (20.95 percent), and argon (0.93 percent). Water vapor is a variable constituent of the atmosphere and ranges from a small fraction of a percentage point to several percentage points. The concentration of carbon dioxide is about 0.038 percent, or 380 parts per million. The concentration of methane is about 1.7 parts per million. The nitrous oxide concentration is about 0.31 part per million. Ozone is a variable constituent of the troposphere and ranges from about 0.02 to 10 parts per million. The atmospheric concentrations of CFC-12 and CFC-11 are only 0.00038 and 0.00023 part per million, respectively. All these GHGs, with the exception of water vapor, are produced by human activities, such as the burning of fossil fuels and biomass (trees, vegetation, and agricultural stubble), the use of nitrogen fertilizers on agricultural areas, and various industrial activities. Each of the GHGs produced by human activities is increasing in concentration in the atmosphere with time. Carbon dioxide is increasing at a rate of about 0.4 percent per year; methane is increasing at a rate of about 1.1 percent per year; nitrous oxide is increasing at a rate of about 0.3 percent per year; and CFC-11 and CFC-12 are each increasing at a rate of about 5 percent per year.

Estimates suggest that tropospheric ozone is also increasing with time, with increases of between 1 and 2 percent per year over North America and Europe during the 1970's and 1980's. Collectively, methane, nitrous oxide, tropospheric ozone, and chlorofluorocarbons are now estimated to trap about as much infra-



The greenhouse effect is aptly named: Some heat from the Sun is reflected back into space (small squiggled arrows), but some becomes trapped by Earth's atmosphere and re-radiates toward Earth (straight arrows), heating the planet just as heat is trapped inside a greenhouse.

red radiation as does carbon dioxide alone. These gases absorb infrared radiation in the spectral region from about 7 to 13 microns, known as the atmospheric infrared window, where water vapor and carbon dioxide do not absorb. Water vapor is a strong absorber below about 7 microns, and carbon dioxide is a strong absorber above about 13 microns. If the rates of increase of the GHGs persist with time, the greenhouse effect of the other gases, along with that of carbon dioxide, could amount to an effective doubling of carbon dioxide levels by the middle of the twenty-first century. This is many decades sooner than the level of carbon dioxide alone is likely to double.

A doubling of carbon dioxide in the atmosphere could lead to a global temperature increase of about 1.2 kelvins if there were no other changes in the climate system. However, a warming caused by an increase of carbon dioxide would lead to more evaporation of water vapor from the oceans and would permit the atmosphere to hold more water vapor, as the capacity of air to hold water vapor increases with increasing temperature. Because water vapor is itself a greenhouse gas, as its concentration increases in the atmosphere, the Earth would warm even further. The net result of these different feedback processes could be that for a doubling of atmospheric carbon dioxide, the surface temperature could increase by about 4 kelvins. It is estimated that the average surface temperature has increased globally by about 0.7 kelvin over the twentieth century, although this temperature increase cannot unambiguously be attributed to the buildup of GHGs over this period. Some temperature data sets also indicate that the 1990's was the warmest decade since the mid-1800's. On the other hand, satellite measures of global temperatures show that no warming has occurred in the first decade of the twenty-first century.

The sources of the GHGs, with the exception of water vapor, are mostly human-initiated. Carbon dioxide results from the burning of fossil fuels and of living and dead biomass, such as the burning of rain forests to clear land in the tropics and the burning of agricultural stubble after harvests. Methane is produced as a combustion product of biomass burning, natural gas leakage, and by the action of anaerobic bacteria in wetlands (such as rice paddies), in landfills, and in the stomachs of ruminants (such as cattle and sheep). Nitrous oxide is produced as a combustion product of biomass burning and by the action of nitrifying and denitrifying bacteria (which add and remove nitrogen) in natural, fertilized, and burned soils. Ozone results from atmospheric reactions involving methane, carbon monoxide, and oxides of nitrogen, which are produced by the burning of fossil fuel and by biomass burning. Chlorofluorocarbons (CFC-11 and CFC-12) are released into the atmosphere when they are used as propellants in aerosol spray cans, as blowing agents for foam insulation, and as refrigerants.

There are a number of other important players in the climate scenario whose roles must be more fully understood before scientists can completely assess the impact of increasing concentrations of GHGs on the future climate. Climate factors that must be studied in more detail include the formation and role of clouds, the heat capacity and circulation of the oceans, and the variability of the Sun's luminosity. The Earth's climate is a complex system affected by many processes and parameters. Experts need to learn more about the interplay of atmospheric GHGs, clouds, the ocean, and solar variability before they may accurately assess the future climate.

METHODS OF STUDY

Studies of the greenhouse effect and its impact on the Earth's climate are multidisciplinary and involve theoretical computer modeling, laboratory studies of the spectroscopic parameters and properties of GHGs, atmospheric measurements of GHGs, and aircraft and satellite measurements of parameters that control climate. Theoretical computer models of climate include zero-dimensional, one-dimensional, two-dimensional, and three-dimensional models. Zero-dimensional models give climate parameters that represent an average for the entire system, such as the mean temperature of the Earth's surface. One-dimensional models are used to study climate in either a horizontal (latitude) or vertical (altitude) direction. In these models, a latitude-dependent surface temperature is the climate parameter of major interest. The vertical one-dimensional model is known as a radiative-convective model and is used to study the effects of changes in concentrations of GHGs on the surface temperature.

Two-dimensional models involve characterizing the temperature variation as a function of both latitude and altitude. The most complex climate model is the three-dimensional model, or the general circulation model (GCM). This model gives a complete description of climate as a function of latitude, longitude, and altitude.

Major uncertainties exist in the understanding of several key parameters in the theoretical modeling of climate. A major uncertainty in these models is the role of cloud feedback, particularly as the Earth heats up. Calculations indicate that as the temperature of the Earth increases, clouds will appear in greater quantity. Clouds both reflect incoming solar radiation and trap Earth-emitted thermal infrared radiation. Hence, clouds can enhance or decrease global warming. The radiative properties of clouds and how they will vary with a warmer Earth are poorly understood and therefore are not included in most theoretical climate models. However, a number of studies have found that cloud cover has a negative feedback effect, suggesting that Earth's climate system is much less sensitive than most computer models indicate. If this negative feedback proves true, global warming would produce a temperature increase of less than 1 kelvin by 2100.

To understand the greenhouse effect of gases, scientists design laboratory experiments to measure the spectroscopic parameters and properties of GHGs. These studies provide information on the infrared wavelengths absorbed by these gases and the intensity of the absorption. These laboratory spectroscopic studies involve filling an absorption cell with a known concentration of a greenhouse gas and measuring the absorption of infrared radiation as a function of spectral wavelength using a scanning spectrometer.

To assess the impact of atmospheric GHGs on the Earth's climate, scientists must obtain accurate measurements of those gases that are found at trace levels in the atmosphere—for example, at atmospheric concentrations ranging from parts per million to parts per trillion by volume. Measuring atmospheric GHGs at these very low atmospheric levels involves a series of different analytical chemistry instruments. Carbon dioxide may be measured with a gas chromatograph equipped with a thermal conductivity detector or with nondispersive infrared instrumentation. Methane is measured with a gas chromatograph equipped with a flame ionization detector or with nondispersive infrared instrumentation. Nitrous oxide, CFC-11, and CFC-12 are measured using a gas chromatograph equipped with an electron capture detector. Ozone may be measured using various ultraviolet and infrared absorption techniques. Water vapor is measured using a hygrometer. Aircraft and satellite measurements have provided very important data on several of the parameters that control climate, including the flux of incoming solar radiation, the flux of outgoing Earth-emitted thermal infrared radiation, the flux of solar radiation reflected by clouds, and the geographical and temporal variability of clouds and surface temperature.

CONTEXT

The buildup in the atmosphere of GHGs such as carbon dioxide, methane, nitrous oxide, tropospheric ozone, CFC-11, and CFC-12, could result in a global warming of the Earth. Atmospheric GHGs are increasing with time. These GHGs result from a variety of human activities, including the burning of fossil fuels (carbon dioxide), the burning of living and dead biomass (carbon dioxide, methane, and nitrous oxide), and the application of nitrogen fertilizers and burning of agricultural and grasslands (nitrous oxide). GHGs are also produced from rice paddies, cattle, and sheep (methane), and from several industrial applications (CFC-11 and CFC-12).

A global warming of the Earth from the buildup of GHGs in the atmosphere would have a significant impact on people's daily lives. Most areas would experience more days per year when temperatures exceed 32° Celsius, and the growing seasons and patterns of rainfall would change. One of the most important effects would be a predicted increase in the height of the world's oceans. The increased height of the oceans would result from the thermal expansion of seawater because of the Earth's high temperature (water is a compressible fluid and expands in volume when heated) and because of the melting of polar ice and snow as the Earth becomes warmer. It has been estimated that a global temperature increase of about 4 kelvins might result in a 2-meter increase in the height of the world's oceans. This increase in water level would cause the flooding of low-lying land areas occupied by more than 40 million people worldwide. Higher temperatures resulting from the buildup of atmospheric GHGs would tax the world's air-conditioning facilities, which require the burning of fossil fuels for their operation. Ironically, the burning of fossil fuels is a major source of atmospheric GHGs.

It must be emphasized that there are fundamental uncertainties and deficiencies in scientists' understanding of climate and the processes and parameters that control it. More must be learned about the effect of clouds and oceans on climate and how these phenomena can affect the climate as the Earth begins to warm. Theoretical computer models of climate are not complete; more research is needed before the future climate of the Earth can be assessed with greater certainty.

Joel S. Levine

FURTHER READING

Ahrens, C. Donald. *Essentials of Meteorology*. 5th ed. Florence, Ky.: Brooks/Cole, 2007. A well-written, student-friendly introductory meteorology textbook. Includes a CD-ROM to aid self-instruction.

Andrews, David. *An Introduction to Atmospheric Physics*. Cambridge, England: Cambridge University Press, 2000. A well written college textbook. Includes information on the enhanced greenhouse effect, ozone depletion, and development of weather systems. Can be technical.

Cook, Alex. *The Greenhouse Effect: A Legacy*. Indianapolis: Dog Ear, 2007. A good introductory work for the general audience. Facts are presented through an easy-to-read narrative.

Environmental Protection Agency. "The Greenhouse Effect: How It Can Change Our Lives." *EPA Journal* 15 (January/February, 1989). Popular, nontechnical accounts of the impact of climate change on agriculture, forests, energy demand, and other areas in a special issue devoted to the greenhouse effect. The principles that control and regulate global climate, including the greenhouse effect, are presented simply for a general audience. Well illustrated with photographs and charts.

Frederick, John. *Principles of Atmospheric Science*. Sudbury, Mass.: Jones and Bartlett, 2008. A comprehensive work covering all areas of atmospheric physics, including the greenhouse effect. Geared toward undergraduate students.

Goody, R. M., and J. C. G. Walker. *Atmospheres*. Englewood Cliffs, N.J.: Prentice-Hall, 1972. A good nontechnical description of solar and infrared radiation in planetary atmospheres and how the radiation balance controls the temperature of a planet. Includes sections on the Sun and the planets, solar radiation and chemical change, atmospheric temperatures, winds of global scale, condensation and clouds, and the evolution of atmospheres.

Hansen, Joel E., and T. Takahashi, eds. *Climate Processes and Climate Sensitivity*. Geophysical Monograph 29. Washington, D.C.: American Geophysical Union, 1984. A collection of papers dealing with various aspects of the climate system, including atmosphere and ocean dynamics, the hydrologic cycle and clouds, albedo and radiation processes, polar ice, and ocean chemistry. Each paper was written by an expert in that particular area of climate research. The papers summarize what is known, along with the major uncertainties and deficiencies in understanding of the processes and parameters that control climate.

Henderson-Sellers, A., ed. *Satellite Sensing of a Cloudy Atmosphere: Observing the Third Planet*. London: Taylor and Francis, 1984. Each chapter is by an active researcher in the area covered, including radiation and satellite sensors, the Earth's radiation budget and clouds, water and the photochemistry of the troposphere, vertical temperature sounding of the atmosphere, cloud identification and characterization from satellites, and the remote sensing of land, ocean, and ice from space.

Levine, Joel S., ed. *The Photochemistry of Atmospheres: Earth, the Other Planets, and Comets*. Orlando, Fla.: Academic Press, 1985. A comprehensive textbook covering atmospheric composition, chemistry, and climate, the sources and sinks of GHGs, and the climate modeling of the Earth. The chapter on climate includes discussions of zero-dimensional, one-dimensional, and three-dimensional (global circulation) climate models and the underlying physical, radiative, and dynamic processes and parameters in each model.

National Research Council. *Changing Climate: Report of the Carbon Dioxide Assessment Committee*. Washington, D.C.: National Academy Press, 1983. Technical report that addresses the possible impacts of climate change on sea level, agriculture, plant growth, and society in general. Considers future carbon dioxide emissions from fossil fuels, the dissolution of carbon dioxide into the oceans, the biosphere storage of carbon dioxide, and the impact of increased carbon dioxide on climate, agriculture, and sea level.

See also: Auroras; Earth-Sun Relations; Earth System Science; Earth's Atmosphere; Earth's Oceans; Planetary Atmospheres; Titan; Triton; Van Allen Radiation Belts; Venus's Atmosphere; Venus's Volcanoes.

H

Habitable Zones

Category: Life in the Solar System

Habitable zones are the places beyond Earth where there is the best chance of finding life. As such, they are a major focus of scientific consideration and investigation that inspire exploratory efforts.

OVERVIEW

At root, the idea of a habitable zone around a star is simple. A planet must not be too close to a star to be too hot for life. On the other hand, it must not be too far from the star to be too cold for life. The spherical shell around a star where a planet will be "just right" for life is the habitable zone of that star. Another name is the circumstellar habitable zone; this name distinguishes it from the galactic habitable zone, the region of a galaxy most favorable to life.

Unfortunately, it is not a simple matter to calculate the limits of this shell for a given star. However, using Earth and the Sun as the basic measure and noting that energy concentration of light from a star decreases as the square of the distance, scientists can predict the rough average radius of the habitable zone around a star. The result is simply the square root of the ratio of the stellar brightness to that of the Sun. The radius will then be expressed in astronomic units (AU), where 1 AU is the average distance of Earth from the Sun. The brightness of the star must be rated at a standard distance, known as its bolometric luminosity.

To complicate matters, a planet might not stay in the habitable zone. To be habitable, a planet should have an orbit that remains in the habitable zone for billions of years. Only planets with relatively circular orbits can do that. Also, the habitable zone moves away from a star as that star ages and grows hotter. Stars are not all the same. Some burn their fuel quickly and do not last for billions of years. The habitable zone of a bright, hot star will not be the same as that of a dim, cool one. Thus, the habitable zone is determined by the star, the planet, and the life-form under consideration.

The type of life involved is, of course, a critical consideration. We cannot discuss habitable zones without first establishing the expectations we have of life around a star. We can identify three very basic requirements for life. First, living things have bodies. Second, a life-form uses a flow of nutrients and energy to sustain its body and bodily processes. Third, life reproduces itself. Reproduction requires bodies (and, most likely, molecules) able to retain the complex and detailed information required for constructing more living forms. Life may be more than this, but it will surely never be less.

With these criteria we find, by an argument too extended to give here, that life in a habitable zone will be water- and carbon-chemistry-based. The habitable zone, then, can be calculated based on the requirement that a planet in the zone will be able to hold water in liquid form long-term. The actual calculation is complicated by many factors; chief among them is the fact that water vapor is a major greenhouse gas. Hence, one cannot simply find the incident energy from the star at various distances because the presence of water retains the heat supplied by the star and thereby expands the habitable zone. An early estimate by Michael Hart had the habitable zone of our Sun between 99 and 105 percent of the Earth's current distance. This was too conservative; a more likely estimate is 95 to 137 percent.

The habitable zone is unique to the star and is determined by the stellar mass and, to a lesser extent, the age of the star. In terms of spectral classes, stars such as Earth's Sun (in class G) and some K- and F-class stars can have habitable zones. The range also corresponds to the stellar surface temperature range of a bit less than 4,500 kelvins (K) to a bit above 7,000 kelvins. Our Sun, at 5,777 kelvins, is in the middle of this range. Stars with a mass 20 percent or more greater than that of the Sun (that is, 1.2 M_S) will not have habitable zones, because they emit deadly amounts of ultraviolet radiation (UV) along with their visible and infrared radiation (IR). UV destroys water molecules and, at high intensity, will eventually strip a planet of the water critical for life. About 1 percent of stars are so large that they consume their fuel and die long before life can form. Indeed, all stars larger than 1.5 M_S would turn into red giant stars and swallow up any life-bearing planets around them before intelligent life could appear. Stars with more than ten times the mass of our Sun are so intensely bright that planets cannot form around them, because light creates pressure on anything it strikes. This radiation pressure is usually too small to matter, but for these very large stars it is great enough that all the material around the star that might eventually form planets is pushed away from the star and is disbursed too quickly for planets to form.

On the other hand, stars with less than about 0.80 M_S do not produce enough high-energy UV light to support life on any planet; their UV output is insufficient for important atmospheric effects such as ozone creation. Any planet close enough to a star of less than about 0.65 M_S to receive sufficient heat will be so gripped by the stellar gravity that it will show the star one face, as the Moon does our Earth. If this is not the case, an effect called spin-orbit coupling will almost certainly force the rotation rate of the planet to be almost as slow as the planetary year, thus frying the planet on one slowly changing side while freezing it on the other. Mercury is such a case. It revolves around the Sun in 88 days but rotates once every 58.7 days, exactly two-thirds of the orbital period. Both these effects are due to the fact that no planet is perfectly spherical. In either case, the planetary face toward the star will be too hot for water, the side away from the star too cold.

Another factor in habitability is variability of stellar output. Our Sun has an eleven-year sunspot intensity cycle that causes a variation in solar luminosity of about 0.1 percent. However, 18 Scorpii, an almost identical star in the constellation Scorpius with a mass of 1.03 M_S and a temperature of 5,789 K, has a much greater variability over a 9 to perhaps 13 year cycle. If great enough, this would make its habitable zone move in and out rapidly thereby negating its benefits for any planet in a basically fixed orbit.

KNOWLEDGE GAINED

The idea of a circumstellar habitable zone has stimulated wide-ranging research resulting in a significant extension of our knowledge of planetary systems generally, as well as of our own solar system in particular. A circumstellar habitable zone imposes quite severe limitations on where best to look for life in the universe. Responses to these limitations are likewise limited. One either accepts the limitations, at least tentatively, and looks for suitable planets around only suitable stars, or one must in some way challenge the limitations.

Tentative acceptance of the limitations takes us in the direction of what has become a successful search for exoplanets, planets orbiting stars other than our Sun. The list is large and growing. The primary technique used in this search detects stellar motion due to the stellar reaction to the orbital motion of the planet. Since large planets create more stellar reaction that is more easily detected, this technique is biased toward discovering large planets. It is no surprise then that most of the known exoplanets are large. It is a bit of a surprise that they tend to be relatively close to their stars and, hence, are sometimes called "hot Jupiters." If this trend continues, it may require revisions in the theories of how planetary systems form. Hot Jupiters are not expected to harbor life even if they are in a habitable zone.

Another puzzling result of these searches is that exoplanets seem to prefer highly elliptical orbits compared with those in our solar system. Such orbits are risky in that they may take the planet out of the habitable zone annually. On a more positive note, the work on exoplanets has confirmed that, as expected, planets tend strongly to be found around stars with high metal content. (In this context, "metal" means any element other than hydrogen or helium.)

Challenges to the idea of the circumstellar habitable zone have either been attempts to show there are niches of habitability outside the habitable zone or efforts to extend the habitable zone in size or to more types of stars, especially to red dwarfs. This later direction seems promising in light of the discovery of planets around the red dwarf Gliese 581. One of them, Gliese 581 c, is said to be the smallest planet yet discovered in the habitable zone of another star. That, of course, assumes that a red dwarf has a habitable zone.

Looking for niches of habitability in our solar system—and, hence, potentially elsewhere—offers the possibility of confirmation by direct examination in the not too distant future and is accordingly fairly popular. Thus, attention has become focused on Mars and some large satellites of Jupiter and Saturn.

Mars has received the most attention, as attested by missions such as the Mars Exploration Rover (which began operating on the Martian surface in early 2004). The Cassini-Huygens mission to Saturn (which entered into orbit around Saturn in mid-2004) included flybys of Titan, revealing its liquid methane oceans and dense atmosphere, while the Galileo mission to Jupiter in the late 1990's gathered a great deal of data on two of Jupiter's satellites, volcanic Io and Europa, whose deep ice sheet appears to have water in liquid form beneath it.

CONTEXT

The dream of "other races of men" on other worlds has been the currency of cosmological speculation at least since the ancient Greek atomists. Men on the Moon were described by the ancient Pythagoreans and in the seventeenth century by Johannes Kepler, and even the eighteenth century philosopher Immanuel Kant gave opinions on the inhabitants of Mars, Venus, and Jupiter. Modern science has tried to inform and thereby reduce this speculation. The concept of a habitable zone is a product of this effort, although it imposes limitations that would, no doubt, have disappointed earlier enthusiasts.

Enthusiasm for finding life elsewhere in the universe is by no means dead. The high profile of the Search for Extraterrestrial Intelligence (SETI) and the advent of the new academic discipline of astrobiology are proof of that. Both of these developments are inextricably connected with the concept of habitable zones and are all but inconceivable without it. The prospect of habitable zones has also stimulated thinking and research in other areas. One such development is the idea of a galactic habitable zone.

The concept of habitable zones also connects to larger cosmological issues, such as questions of the "fine tuning" of the universe that makes life possible somewhere in the universe and the related issue of the anthropic principle, the notion that the universe must contain conditions that allow for the existence of an observing intelligent life-form.

John A. Cramer

FURTHER READING

Aczel, Amir D. *Probability 1*. New York: Harcourt Brace & Company, 1998. Aczel argues that the large number of stars outweighs the limitations of habitable zones to the point where intelligent life must occur throughout the universe.

Cohen, Jack, and Ian Stewart. *Evolving the Alien: The Science of Extraterrestrial Life*. London: Ebury Press, 2002. Cohen and Stewart dispute that alien life will be similar enough to terrestrial forms to frame a meaningful idea of a habitable zone. They also argue the case for various niches of habitability.

Cramer, John A. *How Alien Would Aliens Be?* Lincoln, Nebr.: Writers Club Press, 2001. The first half of the book shows how physical constraints limit where intelligent life might be found in the universe. Hence, it surveys many of the limitations that lead to the idea of a habitable zone and the possibility of habitable niches.

Dole, Stephen H. *Habitable Planets for Man*. 2d ed. New York: Elsevier, 1970. Something of a classic on habitable planets, this is one of the earliest discussions of habitable places for human colonization. It gives a good if somewhat dated account of what makes a place habitable for intelligent life, a more restrictive notion than a habitable zone for any life.

Gonzalez, Guillermo, and Jay W. Richards. *The Privileged Planet: How Our Place in the Cosmos Is Designed for Discovery*. Washington, D.C.: Regnery, 2004. Gonzalez and Richards consider the idea that planetary habitability may be connected with the planet's suitability as a platform for observing the universe.

Grinspoon, David. *Lonely Planets: The Natural Philosophy of Alien Life*. New York: HarperCollins, 2004. This is a wide-ranging and readable book covering habitable zones and many related topics.

Ward, Peter, and Donald Brownlee. *Rare Earth: Why Complex Life Is Uncommon in the Universe*. New York: Springer, 2000. Ward and Brownlee make the case that the limitations on habitable zones are severe to the point of making planets like Earth quite rare.

See also: Extrasolar Planets; Extrasolar Planets: Detection Methods; Extraterrestrial Life in the Solar System; Life's Origins; Mars: Possible Life; Red Dwarf Stars; Search for Extraterrestrial Intelligence; Venus's Surface Experiments.

Hertzsprung-Russell Diagram

Category: Scientific Methods

Between 1904 and 1915, Danish astronomer Ejnar Hertzsprung and American astronomer Henry Norris Russell independently discovered significant relationships between the luminosity and surface temperature of stars. The graph or diagram they developed that displays these relationships has become a powerful tool for summarizing many stellar properties and tracing the stages in the "lives" of stars.

OVERVIEW

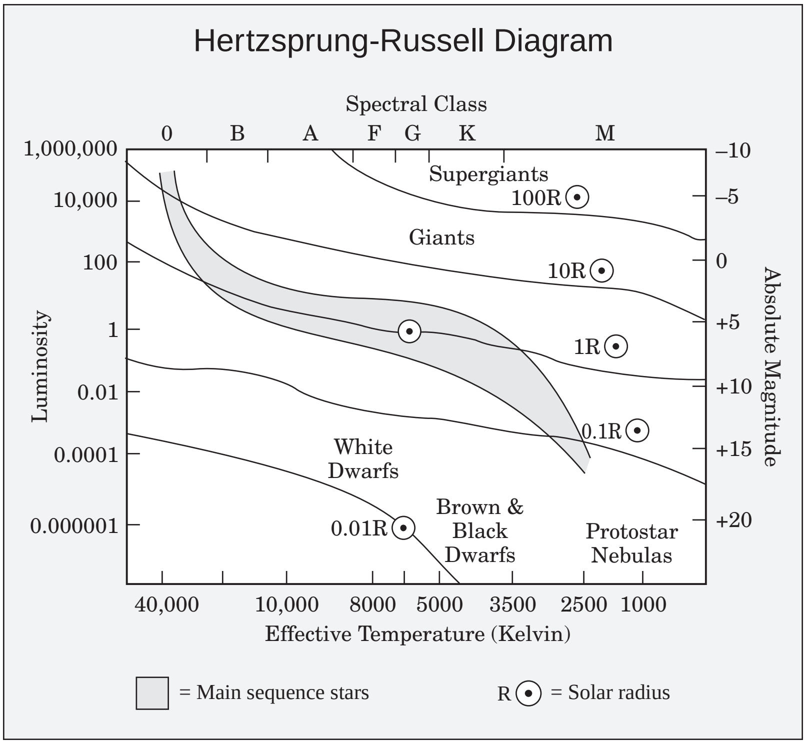
The Hertzsprung-Russell diagram (or simply H-R diagram) is named for the two astronomers, Ejnar Hertzsprung (1873-1967) and Henry Norris Russell (1877-1957), who independently developed a powerful tool that quickly and easily summarizes many properties of a star.

In an H-R diagram, the ordinate (vertical axis) is used to plot some measures of a star's intrinsic brightness; these include absolute visual magnitude, absolute bolometric magnitude, and luminosity. The abscissa (horizontal axis) is used to plot some measures of a star's average photospheric or surface temperature; these include effective temperature, spectral type, and color index. Depending on the particular parameters used to plot stars, the diagram is also known as a color-magnitude diagram or a spectrum-luminosity diagram. However, all variations of the H-R diagram depict the same basic relationships; the slightly different forms of the diagram are due to the specific stellar parameters used in plotting it.

Stars plotted near the top of the H-R diagram are very luminous, while those at the bottom are very dim. Stars plotted on the left side of the diagram have the hottest surfaces and appear bluish, and those on the right side have the coolest surfaces and appear reddish. If spectral type is used, the spectral sequence O B A F G K M is plotted from left to right. The position of a star plotted on an H-R diagram not only indicates the relationship between its luminosity and surface temperature but also provides information about its radius, mass, and stage of development within the stellar life cycle.

The radius is found through application of the Stefan-Boltzmann law for thermal radiating bodies: the rate of emission of electromagnetic radiation per unit surface area is proportional to the temperature (on an absolute scale) raised to the fourth power. Thus the luminosity (L) of a star is proportional to its surface area (A) times its surface temperature (T) raised to the fourth power, or L is proportional to $A \times T^4$. Since the surface area of a sphere is proportional to radius squared, another way of stating this relationship is that luminosity (L) is proportional to radius (R) squared times surface temperature (T) raised to the fourth power, or L is proportional to $R^2 \times T^4$. A star's position in the H-R diagram gives its luminosity and surface temperature, and that information allows one to calculate its radius.

About 90 percent of all known stars fall along a broad band that extends diagonally across the H-R diagram in a "lazy S" shape, from bright,



hot, blue stars at the upper left to faint, cool, red stars at the lower right. This band is known as the "main sequence" and is designated by luminosity class V. The distribution of stars along the main sequence band makes sense intuitively. The stars with the hottest surfaces emit the most light, while those with cooler surfaces emit much less light. Radius does not vary too much along the main sequence, ranging from about five to ten times the Sun's radius at the upper left (the luminous, hot, blue stars) and diminishing to about one-tenth to one-thirtieth the Sun's radius at the lower right (the faint, cool, red stars). It has been found that mass also varies along the main sequence, ranging from several tens to perhaps as much as one hundred times the Sun's mass at the upper left (the luminous, hot, blue stars) and diminishing to about one-tenth to one-fifteenth the Sun's mass at the lower right (the faint, cool, red stars). A star in this band (referred to as "on the main sequence") is in the most stable, longest-lasting chapter of its life. It generates energy by fusing hydrogen into helium in its core. Earth's Sun is a main sequence star with a spectral type and luminosity class of G2 V and an absolute magnitude of about +5, which means it is plotted about midway along the main sequence.

However, not all stars have luminosities and surface temperatures that place them on or near the main sequence. Stars plotted away from the main sequence represent other stages of the stellar life cycle. Some stars are located above and to the right of the main sequence. They are brighter than main sequence stars with the same surface temperature, so by the Stefan-Boltzmann law they must have much larger surface areas from which to radiate, and this means larger radii. Hence these stars are referred to as giants and supergiants. Giants (luminosity class III) typically have about one hundred to ten thousand times the surface area and thus about ten to one hundred times the radius of the Sun. Supergiants (luminosity classes Ia, Iab, and Ib) are even larger, typically having surface areas ten thousand to one million times that of the Sun and radii one hundred to one thousand times that of the Sun. The very largest giants and supergiants are those with the coolest surfaces and hence are red in color. Giants and supergiants have masses ranging from several tens of times the Sun's mass down to about the Sun's mass. They represent stars that have left the main sequence because they exhausted the hydrogen in their cores. An easily seen orange giant (not quite as large or cool as a red giant) is Aldebaran (Alpha Tauri), with spectral type and luminosity class K5 III and absolute magnitude about minus 1, located in the constellation of Taurus the bull; if placed at the center of our solar system, it would fill more than half of Mercury's orbit around the Sun. An easily seen red supergiant is Betelgeuse (Alpha Orionis), spectral type and luminosity class M2 Iab, absolute magnitude about minus 5, located in the constellation of Orion the hunter; if placed at the center of our solar system, it would be about twice as big as the orbit of Mars and extend out into the asteroid belt.

Other stars are plotted below and to the left of the main sequence. They are quite hot but very faint, meaning that they must have small surface areas and radii. They are called white dwarfs and have about the mass of the Sun (about 300,000 Earth masses) packed into a volume about the size of the Earth. They are stars that are near the ends of their lives, having exhausted all their ways of generating energy and shining only because they are still hot.

APPLICATIONS

The Hertzsprung-Russell diagram illustrates the evolutionary stages of a star's life, and it has played a major role in stellar astrophysics. A star is "born" with a certain mass and chemical composition (typically about 75 percent hydrogen and 25 percent helium by mass, with traces of heavier elements). During most of its life, its mass remains nearly constant; only in the last stages of its life does it undergo significant mass loss. On the other hand, its chemical composition, at least in its central region, changes as the star generates energy through a series of nuclear fusion reactions, converting light atomic nuclei into heavier nuclei. A star is in a constant tug-of-war between its tendency to collapse under its own gravity and its tendency to expand due to its pressure. As the star evolves, tapping one energy source after another, its structure changes, and this affects its luminosity and surface temperature, and hence its position in the H-R diagram. By utilizing the H-R diagram, astronomers can easily depict these changes and can trace the evolutionary development of a star as a path followed on the H-R diagram.

Stars are born in interstellar clouds of gas and dust (nebulae) by gravitational contraction. The contraction may be triggered by shock waves from nearby supernova explosions or by an encounter with a galactic spiral arm density wave that compresses the density. As a part of the nebula contracts, it heats up and starts to shine as a protostar. A protostar first appears in the H-R diagram at the extreme right-hand (cool) side. As the protostar continues to contract, it gets hotter and therefore brighter, and it moves diagonally up and to the left in the H-R diagram. When the star's central temperature reaches several million kelvins, hydrogen fusion ignites in the star's core, converting hydrogen to helium. The star stops shrinking and it stabilizes, landing on the main sequence. For a star with the Sun's mass, the protostar stage lasts about 30 million years. More massive stars have stronger self-gravity and contract more rapidly, while low-mass stars have weaker self-gravity and contract more slowly.

A star's location on the main sequence depends on its mass: higher masses toward the upper left and lower masses toward the lower right. The main sequence is the longest, most stable period in a star's energy-producing life. As long as it has hydrogen in its core to fuse into helium, it stays near the main sequence. The length of a star's main sequence stage depends on its mass. Massive stars have more fuel, but they consume it much more rapidly; that is why they are so luminous. Consequently they have very short main sequence lifetimes, no more than a few million years for stars with several tens of solar masses. Earth's Sun is about halfway through its 10-billion-year-long main sequence stage. Low-mass stars have less fuel, but they consume it very slowly; that is why they are so faint. Consequently they have long main sequence lifetimes, much longer than the Sun's and longer even than the present age of the universe; every low-mass main sequence star that ever formed is still a low-mass main sequence star. As a star consumes the hydrogen in its core, it becomes slightly hotter and brighter, and on the H-R diagram, it rises slightly above the zero-age main sequence.

When the hydrogen in the star's core is all transformed into helium, the changes accelerate. The helium core contracts and heats up, and hydrogen fusion is transferred to a shell surrounding the core where it proceeds at a faster rate. The star leaves the main sequence as it becomes more luminous and its outer layers expand and cool. A star like the Sun becomes a red giant. (When that happens to the Sun, in about 5 billion years, it will become about one thousand times brighter than it is now. Earth's oceans will boil away, its atmosphere will escape into space, the rocks of its surface will at least partly melt, and all life on the planet will be extinguished.) Eventually the core of a red giant gets hot enough to fuse helium into carbon. With a nuclear fusion reaction generating energy in the star's core, once again it enters a stable stage, but this phase is brief. A star like the Sun quickly consumes the helium in its core in no more than 1 billion years (compared to 10 billion years for hydrogen fusion in its core while on the main sequence), but it is not massive enough to shrink sufficiently to get hot enough to start any more nuclear fusion reactions.

Instead it puffs off its bloated atmosphere, losing as much as half of its mass. The ejected outer layers are called a planetary nebulae. (A "planetary nebula" has nothing to do with planets; the term developed in the nineteenth century, when these expanding bubbles of gas were seen as round, like planets, and fuzzy, like nebulae, through the telescopes then in use.) Ultraviolet radiation from the hot core often ionizes these gases, making them shine. Examples of planetary nebulae are the Helix nebula in the constellation Aquarius and the Ring nebula in the constellation Lyra. A star's own stellar wind, along with thermal pulses, will drive these outer shells away as expanding bubbles of gas, leaving an exposed hot core as the central star of the planetary nebula.

Such stars are located in a hook-shaped region at the extreme left side of the H-R diagram. The remaining core contracts as much as it can and enters the white dwarf stage. A white dwarf shines only because it is hot, but as it shines it radiates its heat away and becomes cooler and fainter, ending as a cooled-off black dwarf.

A very massive star (more than several times the mass of the Sun) becomes a red supergiant when it leaves the main sequence. It is massive enough for its core to shrink enough to get hot enough to initiate fusion reactions that form elements up to iron. Iron is the heaviest nucleus that can form in fusion reactions that release energy; heavier nuclei require the input of energy to form via fusion. The iron core collapses, sending shock waves through the star, which explodes as a supernova. The outer part of the star is blown away, leaving a collapsed core that becomes a neutron star or in some cases a black hole.

The position of the main sequence in the H-R diagram is affected slightly by the chemical composition of the stars that reach it. As supernovae explode and enrich the interstellar material with elements heavier than helium, new generations of stars will be born with slightly increased abundances of heavier elements, and this shifts the main sequence slightly farther to the right in the H-R diagram.

CONTEXT

Ejnar Hertzsprung's interest in astronomy was fostered by his father, who, although educated as an astronomer at the University of Copenhagen, worked for the Danish Department of Finance. He encouraged his son's interest in astronomy as an avocation, not a vocation, believing that it was not possible to make a living studying the stars. Hertzsprung graduated with a degree in chemical engineering in 1898 from the Polytechnical Institute in Copenhagen and began work in St. Petersburg, Russia. In 1901, Hertzsprung went to Leipzig and spent a year studying photochemistry in Friedrich Wilhelm Ostwald's laboratory. Photography was developing as a serious scientific tool, and Hertzsprung realized its inherent advantages in astronomy, particularly in studying the spectra of stars.

In 1902, Hertzsprung returned to Denmark, where he corresponded regularly with astronomer Karl Schwarzschild. In 1905 and 1907, he published two papers on stellar spectra and magnitudes. In these papers, he pointed out the distinction between red stars that were very luminous and those that were not. He realized that this indicated a significant difference in the size of the stars, and he named them giants and dwarfs respectively. Hertzsprung did not include a diagram with either his 1905 or 1907 paper, but in 1911 he published graphs of the relationship of color to magnitude based on the stars in several star clusters, including the Pleiades. Hertzsprung noted that stars in the diagrams could be divided into two groups: a more populous one, later known as the main sequence, and a smaller group currently recognized as giants and supergiants.

Henry Norris Russell, an American astronomer, began to measure stellar parallaxes using photographic techniques in 1903 in Cambridge, England. By 1910, he had accumulated hundreds of photographic plates. Russell's graphical analysis of the absolute magnitude and spectral type of different stars revealed an interesting correlation. The stars were not scattered randomly over the graph; for most stars, as the luminosity (as measured by absolute magnitude) decreased, so did the surface temperature (as determined from the spectral type). In December of 1913, Russell presented a graph of the relationship (later known the H-R diagram) to the American Astronomical Society. In his address, he also identified giant and white dwarf stars, for which he had laid the theoretical foundations in his papers of 1910 and 1912.

Through the filter of history, priority for the idea for the diagram goes to Hertzsprung. However, Russell was unaware of Hertzsprung's work when he presented his diagram in 1913, based on hundreds of stars he studied from 1903 to 1910, and began to relate the graph to theories of stellar evolution. The origin of identifying the graph as the "Hertzsprung-Russell diagram" is not clear. Hertzsprung himself often remarked that it should be called a color-magnitude diagram for clarification purposes. Calling it the Hertzsprung-Russell (or H-R) diagram seems to have evolved gradually, helped along in its use by the English astronomer Sir Arthur Stanley Eddington, who in 1924 discovered the mass-luminosity relation of main sequence stars. The H-R diagram was included in the articles and lectures of the Danish astronomer Bengt Strömgren during the 1930's, in which he provided an explanation of what the main sequence represents.

The early 1900's was a time of great advances in astronomy. The development of the H-R diagram was one of the more important. The recognition of a correlation between luminosity and surface temperature resulted in a domino effect for stellar research. The H-R diagram has proved to be a versatile astronomical tool. It provides a simple way to represent the structure and depict the evolution of stars.

Richard C. Jones and Richard R. Erickson

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Contains a thorough description of the H-R diagram and its use in the study of stellar evolution, complete with transparent overlays.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Has a good description of the H-R diagram and its use in the study of stellar evolution.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook. Thorough and well written, it includes a good description of the H-R diagram and its use in the study of stellar evolution.

Korn, Katherine G. "Henry Norris Russell (1877-1957)." *Vistas in Astronomy* 12 (1970): 3-6. A general biography of Russell, with insights into his life and work.

Leuschner, A. O. "The Award of the Bruce Gold Medal to Professor Ejnar Hertzsprung." *Publications of the Astronomical Society of the Pacific* 49 (1937): 65-81. A very readable summary of Hertzsprung's astronomical career with an emphasis on the significance of the H-R diagram, at least as it was understood at that time.

Nielsen, Axel V. "Ejnar Hertzsprung: Measurer of Stars." *Sky and Telescope* 35 (January, 1968): 4-6. This obituary notice explores the work and life of Hertzsprung. A good historical look at the man and how his contributions influenced astronomy.

Philip, D. A. G., and L. C. Green. "The H-R Diagram as an Astronomical Tool." *Sky and Telescope* (May, 1978): 395-400. Contains summaries of the International Astronomical Union's Symposium No. 80 of 1977, in which the H-R diagram was examined from various points of view. Describes how the data should be presented to illuminate particular problems, such as the distribution of stars and their chemical composition and age, as well as the variations from one star cluster or galaxy to another.

Russell, Henry Norris. "Relations Between the Spectra and Other Characteristics of the Stars." *Nature* 93 (1914): 227-230, 252-258, 281-286. The paper in which Russell illustrates the relationship between absolute magnitudes and spectral types.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A college textbook for introductory astronomy courses, divided into many short sections on specific topics. Contains a thorough discussion of the H-R diagram, its interpretation, and its use in the study of stellar evolution.

Sitterly, Bancroft W. "Changing Interpretations of the Hertzsprung-Russell Diagram, 1910-1940: A Historical Note." *Vistas in Astronomy* 12 (1970): 357-366. A historical perspective addressing the evolving nature of this astronomical tool.

Struve, Otto. "The Two Fundamental Relations of Stellar Astronomy." *Sky and Telescope* (August, 1949): 250-252, 262. This is a fine historical article on the spectrum-luminosity relationship discovered by Hertzsprung and Russell.

See also: Archaeoastronomy; Coordinate Systems; Earth System Science; Gravity Measurement; Infrared Astronomy; Main Sequence Stars; Neutrino Astronomy; Optical Astronomy; Radio Astronomy; Telescopes: Ground-Based; Telescopes: Space-Based; Ultraviolet Astronomy; X-Ray and Gamma-Ray Astronomy.

I

Iapetus

Categories: Natural Planetary Satellites; The Saturnian System

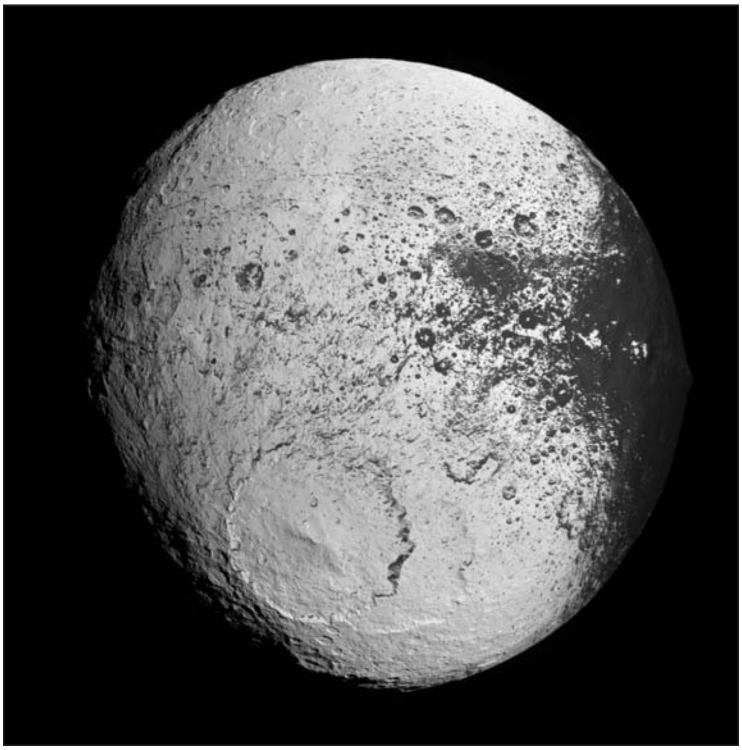
The Saturnian satellite Iapetus is one of the most unusual satellites in the solar system. It has a dark side and a bright side, as well as a ridge along its equator that sits atop an equatorial bulge.

OVERVIEW

Iapetus was first noticed on one side of Saturn by Giovanni Domenico Cassini in October, 1671. Searching for it later on the other side of the planet was futile with the telescope that he was using. He tracked it many times over several years, but only when he had a better telescope in 1705 did Cassini finally see Iapetus on the other side of the planet. He concluded that Iapetus had a dark side and a bright side. It was named for the Titan god Iapetus, a brother of Cronus (the Greek name for the Roman god Saturn) in Greek mythology. Iapetus was originally called Saturn V, because the Saturnian satellites were originally numbered, and many scientists continued to use the numbers. With the discovery of other Saturnian satellites, Iapetus's scientific referent changed to Saturn VII and eventually Saturn VIII. It is one of the sixty satellites of Saturn that had been discovered by 2008. Iapetus's geological features, except for its dark region, are named after characters and places from the French epic poem *The Song of Roland*. The dark region is called Cassini Regio, in honor of Iapetus's discoverer, Cassini.

As Cassini deduced, Iapetus always presents the same face to Saturn, meaning it is synchronous with Saturn. Therefore the time of revolution and the time to circle the planet are identical, 79.32 days. Iapetus is prograde, meaning it turns in the same direction as Saturn. Iapetus's orbit is about 15.5° out of the plane of Saturn's equator; that is, its inclination is 15.5°. The inclination is large enough to cause questions about whether Iapetus was "captured" by Saturn's gravitational field or was generated by the same process that produced Saturn and the other satellites. Iapetus's elliptical orbit (its eccentricity is 0.029) and a synchronous satellite that is 3,561,000 kilometers from the planet also add to the doubt about whether Iapetus was formed by the same process that formed Saturn.

Iapetus is an oblate spheroid with a 1,494-kilometer diameter along the axis pointed at

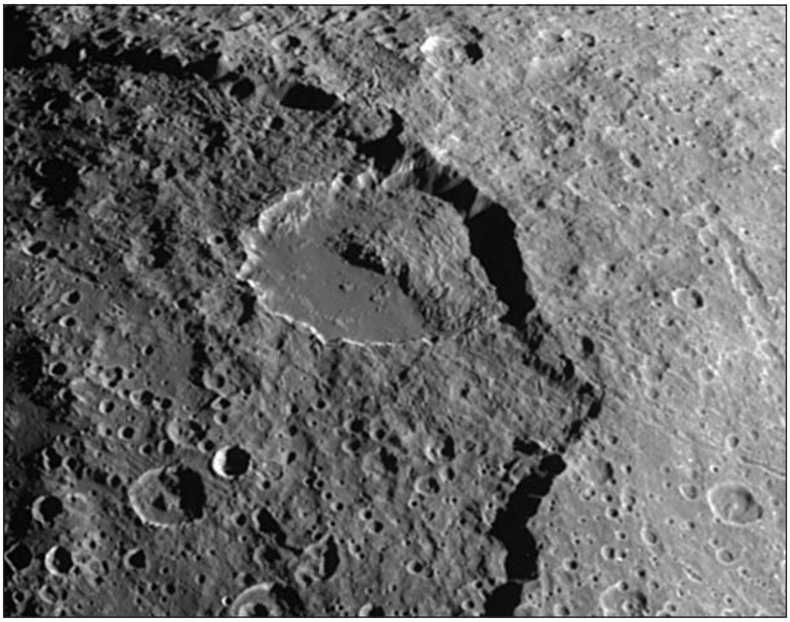


The Cassini spacecraft captured this image of Iapetus in September, 2007. (NASA/JPL/Space Science Institute)

Uranus. The equatorial axis is 1,498 kilometers, and the pole-to-pole axis is 1,426 kilometers. With its equatorial bulge and squashed poles, Iapetus has been said to look like a walnut. A 20-kilometer high ridge that girds most of the satellite along its equator is another striking feature. No other known satellite has a ridge like the one on Iapetus. The ridge is triangular, with a base that is about 200 kilometers wide. The ridge is also cratered, proving it has been in existence for a long time. One theory about the formation of the ridge is that the crust formed while the interior was still flexible enough for the weight of the shell to crush the interior. The interior material forced its way to the surface, fracturing the shell at the equator and forming a ridge. The material cooled relatively quickly, locking it in the new shape that is Iapetus today. This theory requires that Iapetus had to spin much faster during its formation and cooling periods. The slowing of its spin, called despinning, had to take place near a large object such as a planet. This, along with the fact that Iapetus's composition is similar to that of the other Saturnian satellites, indicates that Iapetus was formed in the process that formed Saturn and the other large satellites.

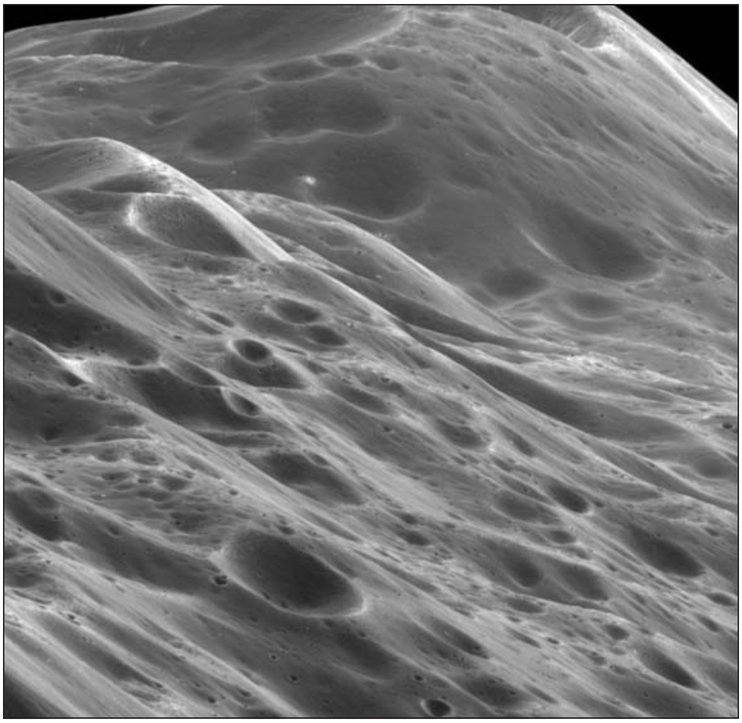
The first feature of Iapetus that astronomers noted was its albedo, which is completely different on the two sides of this satellite. The pictures from the Cassini spacecraft (2007) show a tar-black leading hemisphere with a bright backside hemisphere. The albedo is about 0.05 for the leading side and about 0.6 for the trailing hemisphere. This variation in albedo is noted not only in the visible range but also in the ultraviolet and radio ranges. The Cassini images indicated that the dark material is not solid but is in streaks large enough to appear solid from a distance. Near-infrared spectra indicate that the bright material is ice. The moon's density of 1.1 grams per centimeter cubed indicates that the rock fraction can be no more than 22 percent. The composition of the dark side appears to be ice contaminated with materials such as ammonia, amorphous carbon, poly-hydrogen cyanide (poly-HCN), and hematite (Fe_2O_3).

Some scientists have theorized that the dark material originates from another satellite, Phoebe. Phoebe has a retrograde orbit, meaning that it is traveling in the direction opposite to that of Saturn's rotation and thus in the direction opposite to Iapetus's orbital motion. Material lost by Phoebe because of its retrograde motion can bombard the front face of Iapetus. However, data from the ultraviolet spectra collected by Cassini show that the composition of Phoebe is not the same as the composition of the dark material on Iapetus. There is the possibility that the dark material might be from Hyperion, since the composition of Hyperion does match that of Iapetus. A second option is that the material from Phoebe changes before it bombards Iapetus. The dark material is thought to be a thin layer containing only a small amount (5 percent) of ice. A third opinion is that the dark material is consistent with an external impact. This would cause the poles of Iapetus to be bright, which they



A landslide is seen in the Cassini Regio portion of Iapetus, imaged by the Cassini orbiter in 2004. (NASA/JPL/Space Science Institute) are. The bright poles may be bright because of frost. The peak temperature of the dark side of Iapetus was measured at 130 kelvins, warm enough to allow ice to sublime and refreeze at the poles, producing bright areas. The dark material has a reddish color that might indicate organic compounds. Organics would darken with exposure to radiation. Indications are that the material is porous and in the form of fine particles.

The topography of Iapetus was not well known before the Cassini mission. Even pictures from the Voyager 1 mission did not show any feature on the dark side. The bright side was also largely unknown. Iapetus is cratered, but is not saturated with craters. Around the craters were tall, steep, wall-like features called scarps. It appears that the craters on Iapetus may not have retained the full height of the crater rim, which would indicate the lithosphere is not thick or totally solid. The largest crater is 800 kilometers across with a rim topography of 10 kilometers (that is, raised 10 kilometers from the crater floor). Some scientists believe that Iapetus was formed and its shell hardened at a very early time. The surface might be the oldest surface known.



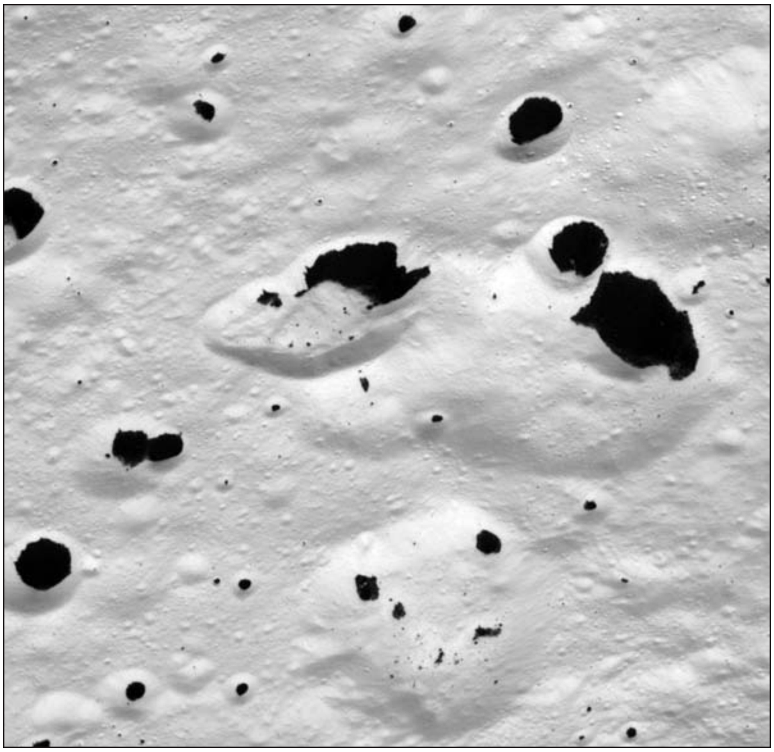
Iapetus's mountainous terrain rises about 10 kilometers above the surface in this Cassini spacecraft image of the moon's equatorial ridge, taken on September 10, 2007. (NASA/JPL/Space Science Institute)

KNOWLEDGE GAINED

Iapetus has been studied with land-based telescopes since it was discovered. Its dichotomous albedo made it an unusual object, and therefore an object of interest. Reflectance spectra taken using the McDonald Observatory produced data that showed reddening of the dark-side material, as dark-side reflectance was compared to that of the bright side. Other land-based studies determined the size of Iapetus, the albedo range, and the longitudinal symmetry of the dark material. One of the most in-depth studies of Iapetus's composition was done from the observatory at Hawaii's Mauna Kea, using both visible and near-infrared spectrometry.

The Voyager 1 mission revealed the diameter of Iapetus. Images of Iapetus were taken and revealed the great amount of cratering that exists on the bright side and at the poles. Perturbations in the path of Voyager allowed scientists to calculate the satellite's density.

In addition to the Image Science Substation (ISS), a version of which was on both Voyagers, the Cassini spacecraft (2004) had an ultraviolet imaging spectrograph (UVIS), which produces simultaneous spectral and spatial images. The visual and infrared mapping spectrometer (VIMS) was used to study the icy satellites by generating reflectance spectra and phase curves, as well as visual pictures. These data are especially important to compare with Earth-based data, providing an evaluation of the Earth-based data and possibly ideas of methods to correct the Earth-based data for two of the problems that occur using Earth-based instruments. Those two problems are the small phase angle seen from Earth and the extra light



Iapetus's frozen surface, splotched with dark terrain, in an image captured by the Cassini spacecraft on September 10, 2007, at an altitude of about 6,030 kilometers. (NASA/JPL/Space Science Institute)

generated by reflection from the rings and from Saturn. The composite infrared spectrometer (CIRS) on Cassini studied the thermal infrared spectrum for emissivity features.

Certain compounds emit a thermal signature in the infrared range that is noticeable in the background thermal spectra. The thermal spectra from Iapetus did not show any strong features. The lack of emission features and the data from near-infrared spectra caused scientists to believe that the surface must be covered with small particles that have a high porosity. One bit of information gleaned from Cassini images is that scarps and crater walls are bright on their north sides and covered with dark material on their south sides.

CONTEXT

Iapetus is certainly an enigma. The black face on one side and the bright face on the other side remain unexplained. Scientists have ideas for explaining these phenomena, but none that seems to offer a complete answer. If the dichotomy is due to material from other satellites, exactly what is the transport mechanism and what changes to the material occur during transport? A second question about Iapetus concerns its formation. How did the equatorial ridge form and why? What does the existence of the ridge tell scientists about the formation of the solar system? Is the surface of Iapetus the oldest surface in the solar system? Can Iapetus lead scientists to determine how the solar system was formed?

Many scientists are surprised that Iapetus can have such a large angle of inclination and be so far away from Saturn, yet still be in synchronous relationship with the planet. Being synchronous requires a very slow revolution, and such slow revolutions are unusual. Iapetus's revolution can be explained only if one concludes that the gravitational force of Saturn has caused despinning. The gravitational force on Iapetus increases and decreases as Iapetus gets closer and then moves away from the planet, and this change in gravitational force causes tidal flexing in the satellite. The heat from this tidal flexing may explain the heat needed to generate some of the satellite's features, such as the ridge and the dichotomous albedo.

C. Alton Hassell and David G. Fisher

FURTHER READING

Bond, Peter. *Distant Worlds: Milestones in Planetary Exploration*. New York: Copernicus Books, 2007. The author discusses several systems, including the Saturnian system and its parts: planet, satellites, and rings. Also discusses how various interplanetary missions have developed our knowledge of the systems from which they have gathered data. Illustrations, bibliography, appendix, index.

Castillo-Rogez, J. C., et al. "Iapetus' Geophysics: Rotation Rate, Shape, and Equatorial Ridge." *Icarus* 190, no. 1 (September, 2007): 179. Using the better data from Cassini, the authors describe several models for the formation of Iapetus. Each model is then evaluated against the known data. Illustrations, bibliography, index.

Corfield, Richard. *Lives of the Planets*. New York: Basic Books, 2007. The author takes the reader through the different space missions. Organized by planet. Index.

Fendrix, Amanda R., and Candice J. Hansen. "The Albedo Dichotomy of Iapetus Measured at UV Wavelengths." *Icarus* 193, no. 2 (February, 2008): 344. The dual face of Iapetus is discussed in view of measurements taken in the ultraviolet range. The meaning of a UV dichotomy is discussed. Illustrations, bibliography, index.

Hartmann, William K., and Ron Miller. *The Grand Tour: A Traveler's Guide to the Solar System*. 3d ed. New York: Workman, 2005. Each major planet, then the major moons including Iapetus, is discussed. Outstanding illustrations; bibliography, index.

McFadden, Lucy-Ann Adams, Paul Robest Weissman, and T. V. Johnson, eds. *Encyclopedia of the Solar System*. San Diego: Academic Press, 2007. The editors have collected articles written by many experts in one of the best scholarly reference works about the solar system. Illustrations, appendix, index.

Thomas, P. C., et al. "Shapes of the Saturnian Icy Satellites and Their Significance." *Icarus* 190, no. 2 (October, 2007): 573. Discusses how measurements from the Cassini mission reveal the shapes of the Saturnian satellites. Illustrations, bibliography, index.

See also: Callisto; Enceladus; Eris and Dysnomia; Europa; Ganymede; Io; Jupiter's Satellites; Lunar Craters; Lunar History; Lunar Interior; Lunar Maria; Lunar Regolith Samples; Lunar Rocks; Lunar Surface Experiments; Mars's Satellites; Miranda; Neptune's Satellites; Planetary Orbits: Couplings and Resonances; Planetary Satellites; Pluto and Charon; Saturn's Satellites; Titan; Triton; Uranus's Satellites.

Impact Cratering

Category: Planets and Planetology

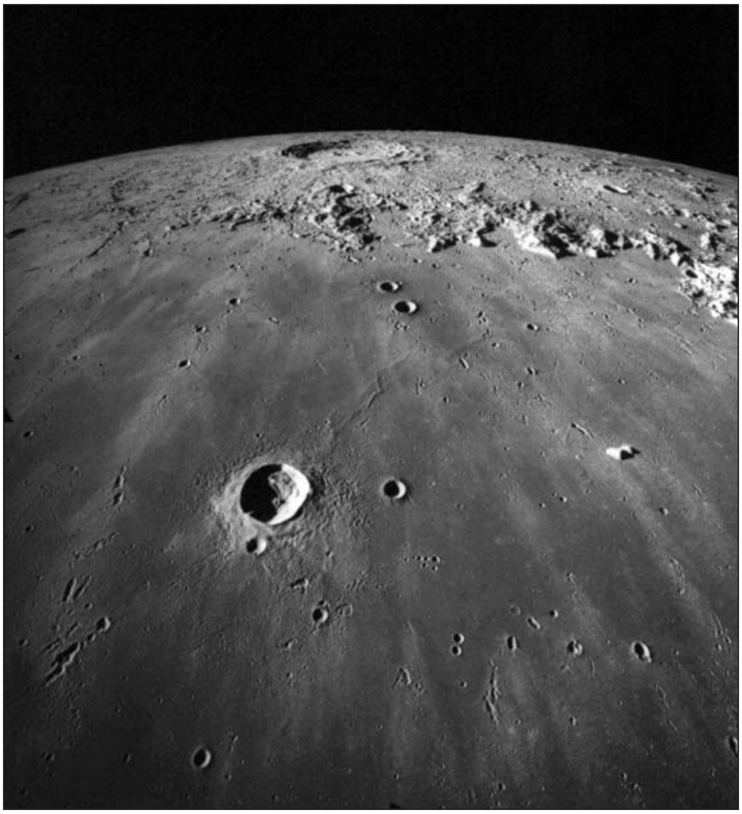
Space-age discoveries about the surface character of other terrestrial planets and satellites around planets throughout the solar system reveal that the early Earth must have been heavily scarred by impacts with planetesimals and minor bodies. Erosion processes and plate tectonics have obliterated most of these ancient craters, but new evidence that major impacts may have had a significant role in shaping the evolution of life has spurred a search for large impact craters.

OVERVIEW

Impact cratering is one of the most fundamental geologic processes in the solar system. Craters have been found on the surfaces of all the solid planets and natural satellites thus far investigated by spacecraft. Mercury and the Moon, bodies whose ancient surfaces have not been reworked by subsequent geologic processes, preserve a vivid record of the role that impact cratering has played in the past. It is inconceivable that the Earth somehow escaped the bombardment that caused such widespread scarring or that it does not continue to be a target for planetesimals still roaming the solar system.

As recently as the 1960's, only a handful of sites on Earth were accepted to be of impact origin. In the early 2000's, the number of confirmed astroblemes (circular surface features considered to have been large impact craters) was well in excess of one hundred and increasing at the rate of several per year. In addition, many "probable" and "possible" impact features are under study. Nevertheless, an enormous discrepancy exists between the number of identified or suspected impact sites on Earth and the number that might be expected.

It is assumed that the flux of incoming bodies is the same for Earth as it is for the Moon. Making allowances for the fact that the Earth is the largest "target" of any of the terrestrial planets and that more than two-thirds of its surface is covered by water, planetologists calcu-



A view of the Moon's Mare Imbrium, taken by the Apollo 17 crew in 1972, shows the cratered lunar surface. (NASA)

late that land areas of the Earth should have been scarred by at least fifteen hundred craters 10 kilometers or more in diameter. In actuality, only about half of the known astroblemes are in this size range. On a global scale, 99 percent of the predicted large impact craters seem to be missing. However, this statistic is not a valid indicator of the impact history of the Earth. Although the impact phenomenon is a geographic process, the probability for discovering impact sites is strongly modified by the geologic stability of various regions of the Earth and by the intensity of the search programs in those areas. Roughly one-half of all the confirmed astroblemes have been found in Canada, which constitutes only one percent of the Earth's surface. In part, this is owing to the stability of Precambrian rock of the Canadian Shield, which thus preserves more of the crater's features, but it also reflects a diligent research effort by Canada's Department of Energy, Mines, and Resources. In general, the number of large impact sites found in the well-explored areas of the Earth agrees with the accepted rate of crater formation on the other terrestrial planets in the past two billion years.

The obvious difference between the surface appearances of Earth and the Moon is explained not by any difference in the rate at which impact craters have formed but in the rate at which they are destroyed. Most of the tremendous numbers of craters on the Moon are more than 3.9 billion years old, while the Earth's oldest surviving astroblemes were formed less than 2 billion years ago. Studies have shown that erosion effectively removes all traces of a 100-meter (diameter) crater in only a few thousand years, and that a 1-kilometer-wide crater, such as the well-known Barringer meteor crater in Arizona, will disappear within a million years. Only craters with diameters greater than 100 kilometers can be expected to leave any trace after a billion years of erosion. This explains not only the absence of widespread cratering on Earth's landscape but also the fact that, among the astroblemes known to exist, medium and large scars are more common than small ones.

Significant craters can be produced only by objects having masses of hundreds of thousands to billions of tons. The Barringer crater, 1.2 kilometers wide and 200 meters deep, is believed to have been formed by a one-million-ton planetesimal that was perhaps 50 meters in diameter. A 27-kilometer-wide astrobleme known as the Nördlinger Ries crater in Germany required an impacting body greater than 1 kilometer in diameter with a mass in excess of 1 billion tons. Planetesimals as large as these two examples are not characteristic of the vagrant meteors that wander through the solar system and occasionally streak into the Earth's skies as shooting stars.

Most of the past impacts on Earth and the Moon appear to be attributable to a family of asteroids known as the Apollo-Amor group (after two specific members of the family). Members of this group are in orbits that graze Earth's orbit and become subject to orbital perturbations that lead them across Earth's path periodically. It is estimated that the average Apollo-Amor object intersects Earth's orbit once every five thousand years, although usually the planet is at some other point on its orbit when this happens. The probability of a collision between Earth and any given Apollo-Amor object is small, but several studies have shown that this family contains between 750 and 1,000 asteroids larger than 1 kilometer in diameter. Statistical analysis suggests that such sizable bodies must collide with the Earth an average of once every 600,000 years.

Impact events involve tremendous transfers of energy from the incoming planetesimal to Earth's surface. A projectile's energy of motion increases only linearly with its mass but as the square of its velocity, so surprisingly large craters result from relatively small bodies traveling at hypervelocities. Depending on the directions of motion of Earth and of the planetesimal, impacts on the planet may involve relative velocities as high as 50 kilometers per second. At velocities surpassing 4 kilometers per second, the energy of the shock wave created by the impact is far greater than the strength of molecular adhesion for either the planetesimal or Earth. Therefore, on impact the planetesimal acquires the properties of a highly compressed gas and explodes with a force equivalent to a similar mass of blasting powder.

The shock wave from this explosion intensely compresses the target material and causes it to be severely deformed, melted, or even vaporized. In all but the smaller impacts, the entire projectile is also vaporized. The shock wave swiftly expands in a radial fashion, pulverizing the target material and intensely altering the nature of the target rock by extreme and almost instantaneous heat and pressure. This is immediately followed by decompression and what is called a rarefaction wave that restores the ambient pressure. The rarefaction wave moves only over free surfaces, so it travels outward over the ground surface and into the atmosphere above the impact. It becomes the excavating force that lifts vast quantities of the pulverized target material upward and outward to create the crater cavity.

The rarefaction wave excavates a hole whose depth is one-third of its diameter and whose profile follows a parabolic curve, but this depression is short-lived and is therefore called the transient cavity. After passage of the rarefaction wave, a large amount of pulverized target material from the walls of the transient cavity slumps inward under gravity, and some of the ejecta lofted straight up into the atmosphere falls back into the excavation. Together, these sources contribute to a lens-shaped region of breccia that fills the true crater's floor and leaves a shallower, flat-floored apparent crater as the visible scar of the impact. Apparent craters generally exhibit a depth of only one-tenth to one-twentieth of their diameters. Meanwhile,

Earth's Largest Impact Craters

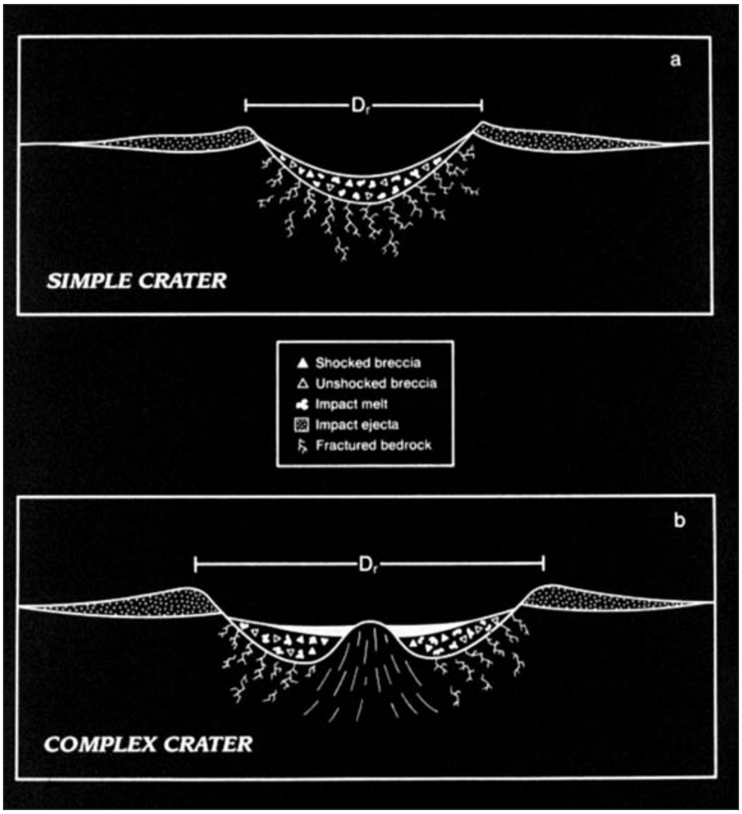
Diameter (km)	*Location*	*Crater Name*	*Age (millions of years)*
300	South Africa	Vredefort	2,023
250	Canada	Sudbury	1,850
170	Mexico	Chicxulub	65
100	Canada	Manicougan	214
100	Russia	Popigai	35
90	Australia	Acraman	590
90	United States	Chesapeake Bay	36
80	Russia	Puchezh-Katunki	175
70	South Africa	Morokweng	145
65	Russia	Kara	73
60	United States	Beaverhead	600
55	Australia	Tookoonooka	128
54	Canada	Charlevoix	357
52	Sweden	Siljan	368
52	Tajikstan	Kara-Kul	5

Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.

the rarefaction wave carries ejecta particles outward over the surrounding landscape, where they fall to Earth as a blanket of regolith that is distinguishable from the local target rock by the effects of shock metamorphism.

METHODS OF STUDY

Impact phenomena are rare enough on the human timescale that no crater-forming events are known to have occurred in recorded history. Owing to this passage of time and to the fact that most existing astroblemes have been severely altered by erosion, impact cratering has been studied by the unique modifications that a powerful impact shock makes in the rocks and minerals at the site. Scientists study the deformation and structural damage to buried strata, and by looking for the presence of certain rare elements and minerals in the sediments surrounding suspected impact sites.



These diagrams show cross sections of the structures of both simple and complex craters. (NASA)

Much attention has been given to the effects of the shock wave on terrestrial rocks, since shock metamorphism is considered to be the most enduring and positive identifier of ancient astroblemes. Shock metamorphism differs from endogenic metamorphism by the scales of pressure and temperature involved and by the very short duration of the exposure to those pressures and temperatures. Endogenic metamorphism usually involves pressures of less than 1 gigapascal (100,000 atmospheres) and temperatures not greater than 1,250 kelvins. The pressures involved in shock metamorphism are exponentially greater, reaching several hundred gigapascals for an instant in the vicinity of the impact. Rock exposed to pressures in excess of 80 gigapascals and temperatures of several thousand kelvins is immediately vaporized. Lesser pressures and temperatures at increased distances from the point of impact produce signs of melting, thermal decomposition, phase transitions, and plastic deformation.

Pockets of melt glass up to several meters thick are commonly found in the breccia within the crater, indicating that pressures there reached 45-60 gigapascals. Coesite and its denser relative, stishovite, are forms of quartz that occur naturally only at impact sites. Shatter cones, conically shaped crystals created at pressures of from 2 to 25 gigapascals, are another prominent feature of shock metamorphism and are particularly well developed in fine-grained isotropic rock. Microscopic examination of impact-shocked porous rock reveals that quartz grains are deformed so as to fill the pores and interlock like the pieces of a jigsaw puzzle. Even at a considerable distance from the impact point, quartz grains tend to be elongated in the direction of the shock wave's passage.

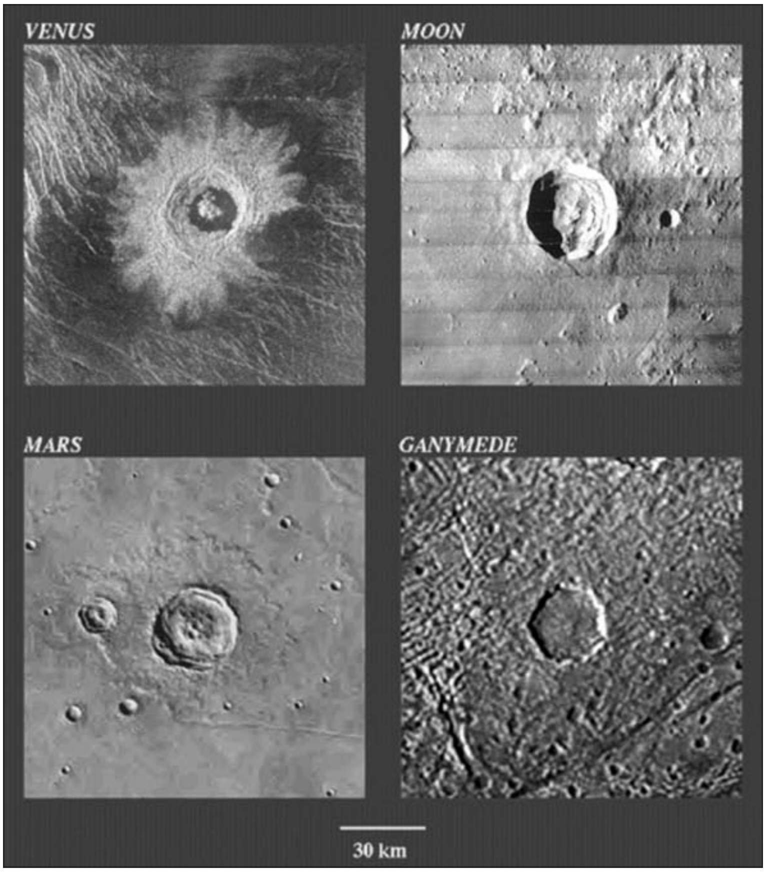
Theories concerning cratering dynamics can also be tested by analogy to some of the craters produced by the detonation of nuclear devices.

This latter technique has adequately explained the morphology of the smaller astroblemes, those with diameters that do not exceed 2-4 kilometers. Larger impact events involve additional dynamics that are not mimicked by nuclear devices thus far tested. Astroblemes greater than 2 kilometers in diameter in sedimentary rock or 4 kilometers in diameter in crystalline rock display a pronounced central uplift owing to an intense vertical displacement of the strata under the center of the impact. An additional feature distinguishing complex craters is that their depths are always a much smaller fraction of their diameters than is the case with simple craters.

Photographic imaging of Earth from space has revealed some young and well-preserved astroblemes in remote and poorly explored areas of Earth, such as the Sahara Desert. More important has been the satellite's ability to reveal structures that still preserve a faint but distinct circularity when seen from orbit, although at ground level they are so eroded that their circularity has escaped detection. One of the largest astroblemes yet discovered was detected from Landsat satellite images in this way. New imaging technologies, including advanced radar and sonar mapping, promise to extend the capabilities of space surveillance and remote sensing in recognizing possible impact sites.

CONTEXT

The degree to which the Earth is in danger of being struck by a massive planetesimal began to be appreciated about the middle of the twentieth century. In 1980, a team led by Nobel Prize-winning physicist Luis Alvarez announced dramatic evidence suggesting that an asteroid impact that occurred 65 million years ago created such planetary stress that it explained a mysterious massive extinction of life-forms known to have occurred on the Earth at that time. At sites all around the world, researchers discovered that clay deposits at the boundary layer between the Cretaceous and Tertiary periods contained up to one hundred times the normal abundance of the metal iridium, which is rare in Earth's crustal rocks but 1,000 to 10,000 times more abundant in the makeup of many asteroids. This Cretaceous-Tertiary boundary layer is coincident with the point at which fully 70 percent of the life-forms then existing on the Earth, including the dinosaurs, became extinct. Further study has also revealed that this same sediment layer is rich in shock-metamorphosed quartz grains, known only to occur naturally from impact explosions.



Examples of impact craters on Venus, Earth's moon, Mars, and Ganymede. (Venus and Moon: Robert Herrick/Lunar and Planetary Institute; Mars: Calvin Hamilton/Los Alamos National Laboratory; Ganymede: Paul Schenk/Lunar and Planetary Institute)

Debate continues as to whether an asteroid impact was the primary cause of the mass extinctions at the close of the Cretaceous period or merely the final factor, but there is general agreement that a colossal impact occurred at that time. The volume of material represented in the boundary sediments suggests that the planetesimal was perhaps 10 kilometers in diameter and would have created a crater of as much as 200 kilometers in width. An astrobleme in the Gulf of Mexico near Belize, called the Chicxulub Crater, closely fulfills these criteria. Many scientists accept it as the impact site for the K-T (German for Cretaceous-Tertiary) event. Meanwhile, several other iridium spikes (abnormally high concentrations of the metal) have been found in the sedimentary beds coinciding with other recognized mass extinctions.

Early in the twenty-first century several researchers put forward candidate craters to mark an impact event dated to the time of what paleontologists often call the Great Dying. At the boundary between the Permian and the Triassic (the P-T boundary), which also marks the end of the Paleozoic and the start of the Mesozoic era, life on Earth was very nearly exterminated. A conservative estimate is that 95 percent of all species died out at that time. Life rebounded and the dinosaurs went on to rule the Earth, until they too were wiped out catastrophically. Of the various craters proposed to have resulted from a P-T boundary impact event, the one that appears most likely to turn out to be correct (if any of them are correct) is a crater located in Antarctica, buried unfortunately under 1.5 kilometers of ice. What provides extra confidence that this crater could be the result of a P-T boundary impact event is the fact that at its antipode is located the Siberian Traps. Energy from the impact would have undergone antipodal focusing through the Earth's core to ravage the area on the planet 180° away from the impact site. The Siberian Traps experienced tremendous amounts of volcanic activity around 248 million years ago, the very time of the P-T boundary and the Great Dying. This scenario remains controversial but, if true, would represent an even larger impact event than the accepted K-T boundary event that gave rise to the Chicxulub Crater.

Three related discoveries suggest the possibility that impact cratering may not be an entirely random process, so far as its distribution through time is concerned. Paleontologists David Raup and J. John Sepkoski, Jr. have shown evidence, based on a rigorous analysis of the marine fossil record, that mass extinctions appear to occur with regularity every 26 million years. Independently, the team of Walter Alvarez (also a member of the team that discovered the K-T iridium anomaly) and Richard Muller have discovered evidence that the ages of major known terrestrial astroblemes seem to be periodically distributed at intervals of roughly 28 million years. For some time, researchers have sought a mechanism that could account for the numerous polarity reversals in Earth's magnetic field over geologic history, and some have suggested that major impact events may be the cause. Several studies have reported an apparent fine-scale periodicity in Earth's magnetic field reversals with a cycle of 30 million years. Although the intervals are not in perfect agreement, they are very close, considering the difficulty of precisely dating extinctions and the exact ages of astroblemes.

These discoveries suggest that there may be an as yet undiscovered member of the solar system that moves in such a way as periodically to disrupt the Oort Cloud, the cloud of comets believed to exist on the fringes of the solar system, causing a barrage of planetesimals to descend upon the inner planets. Although the existence and location of such a body remain speculative and controversial, it has been characterized as a dwarf companion star of the Sun and is called Nemesis.

Richard S. Knapp

FURTHER READING

Consolmagno, Guy. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994. A text accessible to college-level science and nonscience majors alike. Presents most topics using low-level mathematics; involves integral calculus where required. Demonstrates how the area of planetary science progresses by questioning previous understandings in the light of new observations.

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001. A challenging and thorough text for students of planetary geology. Covers extrasolar planets and provides an in-depth, contemporary explanation of solar-system formation and evolution. An excellent reference for the most serious reader with a strong science background.

Dixon, Dougal. *The Practical Geologist: The Introductory Guide to the Basics of Geology and to Collecting and Identifying Rocks*. New York: Fireside, 1992. A beginner's guide to the physical processes that formed the Earth and modified its surface over geologic time. Heavily illustrated with guides for rock and mineral identification.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions of 2003 that have investigated all planets. Takes an astrophysical approach to place the solar system in a wider context, as just one member of similar systems throughout the universe.

Faure, Gunter, and Teresa M. Mensing. *Introduction to Planetary Science: The Geological Perspective*. New York: Springer, 2007. Designed for college students majoring in Earth sciences, this textbook provides an application of general principles and subject material to bodies throughout the solar system. Excellent for learning comparative planetology.

Grieve, Richard A. F. "Terrestrial Impact Structures." *Annual Review of Earth and Planetary Sciences* 15 (May, 1987): 245-270. A thorough summary of what is known about the cratering process on the Earth, written by a leading authority on the subject. It is intended for the scientific reader, but its illustrations, extensive bibliography, and introductory and summary sections are of value even to those who are not familiar with the concepts and terminology in the body of the article.

Hartmann, William K. "Cratering in the Solar System." *Scientific American* 236 (January, 1977): 84-99. Dated, but a comprehensive explanation of the role attributed to impact cratering in shaping the surfaces of all of the terrestrial planets. The author explains the basis for estimating the frequency of impacts for various sizes of planetesimals and the logic behind using crater counts to estimate the ages of planetary surfaces. The article also explains the theory that the first half billion years of solar-system history involved an extremely heavy bombardment of all the inner planets.

_______. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science. Includes chapters on all planets and their systems. Discusses the role of impact cratering in shaping planets and their satellites.

Kerr, Richard A. "When Disaster Rains Down from the Sky." *Science* 206 (November 16, 1979): 803-804. Taking a descriptive approach easily comprehended by laypersons, this article summarizes research by several investigators attempting to compute the frequency with which the Earth is struck by crater-forming bodies. The article places particular emphasis on the Apollo asteroid group and examines suggestions that the Apollo family is supplied with new planetesimals by the decay of former comets.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. Planetary atmospheres are treated as important physical features of the various members of the Sun's family. They are discussed individually in the context of what is known about each planet's characteristics and with regard to theories about their evolution and the evolution of the entire solar system. Geared for the undergraduate college student.

Muller, Richard. *Nemesis: The Death Star*. New York: Weidenfeld & Nicolson, 1988. Despite its tabloid title, this is an excellent discussion of the chain of discoveries leading to the Nemesis theory by the Berkeley physicist who developed it. The volume is organized in two parts: The first recaps the evidence for a major impact at the K-T boundary, and the second tells how further research led Muller to postulate the existence of Nemesis. Intended for lay readers, the book gives insight into how the scientific discovery process works, as well as explaining the theory.

Murray, Bruce, Michael C. Malin, and Ronald Greeley. *Earthlike Planets*. San Francisco: W. H. Freeman, 1981. Although terrestrial impact craters are not specifically discussed, the impact mechanics that produce craters are presented here in terms that are suitable for general readers. A somewhat dated but nevertheless excellent discussion of cratering as a ubiquitous aspect of the surfaces of all the inner planets.

Raup, David M. *The Nemesis Affair*. New York: W. W. Norton, 1986. The author is a significant figure in the field of paleontology and has done leading research on the apparent periodicity of extinctions and magnetic reversals. His narrative is a fascinating personal account of the ideas and the individuals who led the scientific community from extreme skepticism to general acceptance that impact "catastrophism" may have played a major role in the Earth's evolution and its lifeforms.

See also: Lunar Craters; Lunar History; Lunar Maria; Mars's Craters; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Venus's Craters.

Infrared Astronomy

Category: Scientific Methods

Infrared astronomy explores the universe by focusing on wavelengths of the electromagnetic spectrum that are longer than those of visible light. This region of the spectrum is useful for studying the process of star formation, for studying objects that are obscured by clouds of interstellar material, and for studying lower-temperature objects that do not radiate as prevalently in the visible portion of the spectrum.

OVERVIEW

Infrared astronomy focuses its study on wavelengths of electromagnetic radiation that are a little longer than those of visible light. The infrared region of the spectrum covers a wide range of wavelengths from waves slightly longer than those of visible light (0.7 micron, or 0.7 millionth of a meter) to those as long as 1,000 microns. The longest infrared wavelengths are about 1 millimeter in length and mark the boundary with the microwave radio spectrum.

Infrared radiation from distant sources is very difficult to detect. The Sun is so close that the infrared radiation it emits can be detected in the form of heat. The Moon also emits easily detected infrared radiation. However, to detect emissions from other stars, planets, nebulae, or galaxies, very sensitive detectors are needed. The shortest infrared waves are known to astronomers as the "photographic infrared" because they are very similar to visible light and can be detected with certain types of photographic emulsions and other types of optical detectors. At longer wavelengths of infrared radiation, objects can be detected that are not visible at optical wavelengths. Nevertheless, at these wavelengths, the detectors used for the photographic infrared are no longer useful.

Modern infrared detectors often use a substance called indium antimonide, which changes its electrical conductivity when exposed to infrared radiation. In order to be effective, however, it must be kept very cold. Solid nitrogen or liquid helium is used to surround the material to bring its temperature from 50 kelvins down to within a few kelvins of absolute zero. Another long-wavelength infrared detector uses a crystal of the semiconducting material germanium that contains traces of the rare metal gallium. This detector must be kept to a temperature only 2 kelvins above absolute zero.

Earth's atmosphere provides advantages as well as disadvantages to infrared astronomy. Some infrared observations can be done during the day as well as at night, allowing infrared detectors to be mounted on large optical telescopes for daytime use. The disadvantage posed by the atmosphere is that water vapor and carbon dioxide absorb certain wavelengths of infrared radiation, making them invisible to astronomers.

The Spitzer Space Telescope views objects in the infrared wavelength range; this artist's conception shows the spaceborne observatory against the sky as it would appear in the infrared, the bright band being the Milky Way. (NASA/JPL-Caltech)

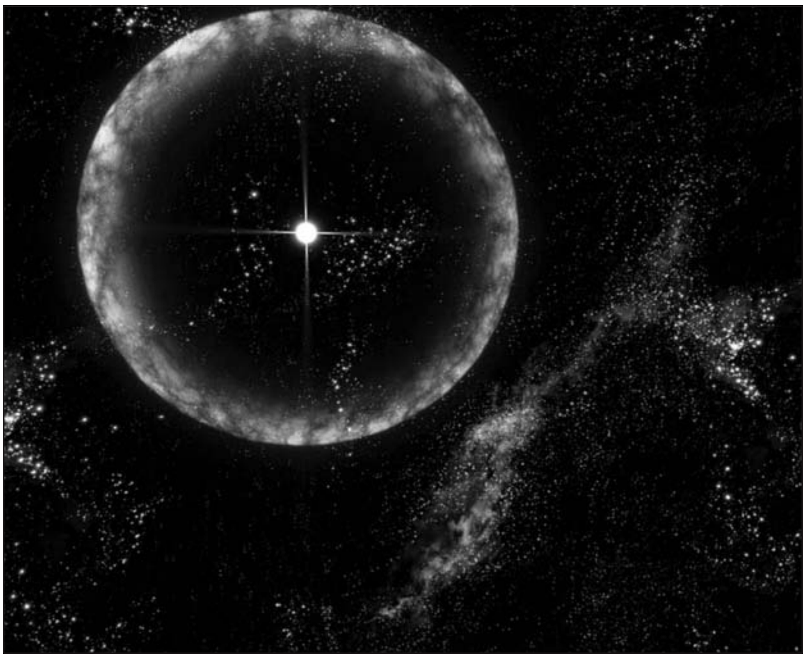
Infrared astronomers therefore have to choose particular wave bands at which the atmosphere allows a clear window. To see through these windows is often a challenge, as common objects—such as telescopes and even the sky—can radiate at these same wavelengths if they are at the appropriate temperature.

Infrared astronomers have designed ways of partially overcoming the problems posed by the atmosphere. Infrared instruments are designed so that no stray radiation from the instrument itself can enter the detectors. To overcome sky brightness in the infrared, astronomers take measurements of the observational target. Measurements include the infrared brightness of both the object and the sky. The telescope is then moved slightly so that it is no longer pointing at the source, where it takes an infrared measurement from the background sky only. When the second measurement is subtracted from the first, it is possible to determine the brightness of the object itself. This technique works well for stars but works less well for objects such as nebulae, which cover a wider field of view. The technique can be modified to scan a wider portion of the sky, making images of larger areas possible. Detectors have been developed that can record such images in a single exposure.

The ideal earthbound infrared observatories are located at very high altitudes and in arid atmospheric conditions. The best site is on Mauna Kea in Hawaii, 4,200 meters above sea level, where two of the world's largest infrared telescopes reside: the National Aeronautics and Space Administration's (NASA's) 3-meter-diameter telescope and the United Kingdom Infrared Telescope (UKIRT), with a 3.8-meter-diameter mirror.

At wavelengths longer than about 30 microns, the atmosphere begins to absorb so much infrared radiation that ground-based observation is impossible. To observe these longer (or far-infrared) wavelengths, astronomers conduct their observations remotely. In the 1970's, ten rocket flights carrying infrared detectors performed a survey of nine-tenths of the sky. During these early flights, it was discovered that the center of the Milky Way and other galaxies were strong sources of far-infrared radiation. High-altitude balloons have also been used to make observations; NASA converted a C-141 transporter plane into a flying infrared observatory, the Kuiper Airborne Observatory, complete with a 0.9-meter telescope. The observatory carries scientists to altitudes of about 12,500 meters, where they can make observations free of about 99 percent of the atmosphere's water vapor.

These types of observations are valuable, but the best way to solve the observational problems posed by the atmosphere is to observe outside the atmosphere completely. Although astronomers have flown many satellites to measure other types of radiation from space, the infrared band has presented difficulties because of the necessity of keeping the detectors at extremely low temperatures. In 1983, a fully dedicated infrared satellite was finally launched. The Infrared Astronomical Satellite (IRAS) was a joint project by the United States, the Netherlands, and England. IRAS investigated the sky from an orbital altitude of 900 kilometers. Throughout its development, this mission proved to be one of the most difficult ever attempted. The infrared detectors had to be designed so that even in orbit they were cooled to within a few degrees of absolute zero with nearly 90.7 kilograms of liquid helium. The lifetime of the satellites was limited because the liquid helium slowly boils away. IRAS was able to function efficiently for a total of ten months. The principal instrument aboard IRAS was an array of sixty-two semiconductors that were sensitive to the majority of the infrared spectrum. The satellite was roughly the size of a small automobile and weighed 1,076 kilograms.



In 2004 this magnetar, a blast of energy from a flare caused by the twisting magnetic field of a neutron star, was so bright that it lit up Earth's upper atmosphere. (NASA)

In spite of the complexity of keeping the instruments cold, the mission was highly successful. IRAS scanned 95 percent of the sky a total of four times at the middle and far-infrared wavelengths. It was able to detect and catalog about 250,000 celestial sources of infrared radiation. IRAS produced such an enormous catalog of infrared sources, significantly expanding the field of infrared astronomy. Numerous space-based detectors and specially dedicated observatories followed. One example of an instrument on an observatory primarily taking data in the visible was the Near Infrared Camera and Multi-Object Spectrometer (NICMOS) installed on the Hubble Space Telescope (HST) by shuttle astronauts in 1997 during the STS-82 servicing mission. Two examples of dedicated observatories that went far beyond the early discoveries of IRAS were the Space Infrared Telescope Facility (later renamed the Spitzer Space Telescope), a member of NASA's Great Observatories program, and ESA's Infrared Space Observatory (ISO).

The European Space Agency (ESA) launched the Infrared Space Observatory (ISO) on November 17, 1995, and placed it in an elliptical orbit ranging from as close to the surface of the Earth as 1,000 kilometers to as high above the surface as 70,500 kilometers; this gave ISO a twenty-four-hour orbital period. ISO was designed to detect infrared radiation ranging from 2.5 to 240 microns in wavelength. To achieve that, ISO was equipped with four separate scientific instruments cooled by liquid helium. The instruments were disturbed by energetic particles when ISO dropped through and rose out of Earth's Van Allen radiation belts, but fortunately the observatory spent 70 percent of each orbit well beyond those disruptive belts. The ISO mission concluded on May 16, 1998, after the cryostat's helium had boiled off, thereby raising the temperature of the instruments sufficiently high to render them useless for infrared measurements.

The Hubble Space Telescope was placed into orbit from the space shuttle Discovery on mission STS-31 on April 25, 1990. One of NASA's four Great Observatories, Hubble was designed to be adaptable and have rotating instruments in its science bay to conduct investigations from part of the infrared through the entire visible portion of the electromagnetic spectrum to portions of the ultraviolet. As first deployed, Hubble suffered from a precise but inaccurate optical prescription for its main mirror. Astronauts on STS-61 in December, 1993, installed a corrective device called Corrective Optics Space Telescope Axial Replacement (COSTAR) which brought incoming light into proper focus, saving Hubble from an otherwise dismal outcome. On a second servicing mission, STS-82, in February, 1997, astronauts removed an instrument from Hubble and replaced it with the Near Infrared Camera and Multi-Object Spectrometer (NICMOS), designed to perform infrared studies involving wavelengths of 0.8 to 2.4 microns.

NICMOS was outfitted with a unique thermal management system making use of a block of solid nitrogen to keep the instruments cooled to about 40 kelvins. Unfortunately, after NICMOS was incorporated into HST in space and afterward put through early commissioning activities and subsequent science observations, it was clear that a heat leak had developed that would raise the solid nitrogen's temperature to a point, after two years or less, where it would no longer provide sufficient cooling for the sensors in the instruments to produce accurate results. Meanwhile, scientists and engineers developed a mechanical cryocooler that would in turn be installed on NICMOS during the next shuttle servicing mission to HST. It was capable of monitoring temperatures between 75 and 86 kelvins, low enough for the instruments to function. In March, 2002, astronauts on space shuttle mission STS-109 saved NICMOS. Further repairs to NICMOS became necessary, and plans were made for another mission to effect those repairs and allow HST to continue front-line infrared observations before its planned replacement by the James E. Webb Space Telescope.

The last member of NASA's Great Observatories, the Space Infrared Telescope Facility (SIRTF), was launched into a solar orbit on an expendable Delta II booster on August 25, 2003 (thirteen years to the day after Hubble was placed in Earth orbit), rather than being deployed from a space shuttle by astronauts. After launch the observatory was renamed the Spitzer Space Telescope (SST) in honor of astronomer and longtime proponent of placing large telescopes in orbit Lyman Spitzer. The Spitzer telescope was designed to observe infrared radiation from 3 to 180 microns for at least five years, being outfitted with an infrared camera (operating from 3 to 180 microns), an infrared spectrometer (operating between 5 to 40 microns), and far-infrared detector arrays. To achieve these goals, the observatory was cooled by liquid helium to 5.5 kelvins, aided greatly by passive cooling with a sunshield. Spitzer was going strong after five years, with much of its liquid helium remaining.

The James E. Webb Space Telescope (JWST) is designed to be set up in an operation position at the L2 Lagrangian point, a spot 1.5 million kilometers from Earth where the gravitational influence on the observatory from Earth will be balanced by that from the Sun. This observatory, fully devoted to infrared astronomy, is designed to detect emissions from 0.6 to 28 microns. JWST's design is such that it will be half as massive as Hubble, but with a folded optical system of 18 individual hexagonal segments with a light collecting area six times greater than HST once fully deployed. To achieve its intended mission and remain cooler than 40 kelvins, JWST is desiged to rely primarily on sunshields. Launch of JWST is targeted for 2013 by the European Space Agency atop an Ariane V booster; JWST will be a NASA observatory with ESA instruments, launched by ESA in exchange for time using the telescope. If successfully set in place at the L2 spot, JWST will begin to fulfill four primary objectives: (1) search for light from the earliest stars and collective structures formed in the first hundreds of millions of years after the big bang; (2) produce images and data that will assist in expanding our understanding of galactic formation and subsequent evolution; (3) produce images and data that will assist in expanding our understanding of planetary formation and subsequent evolution; and (4) search for the existence of organic material essential for the development of life in the universe.

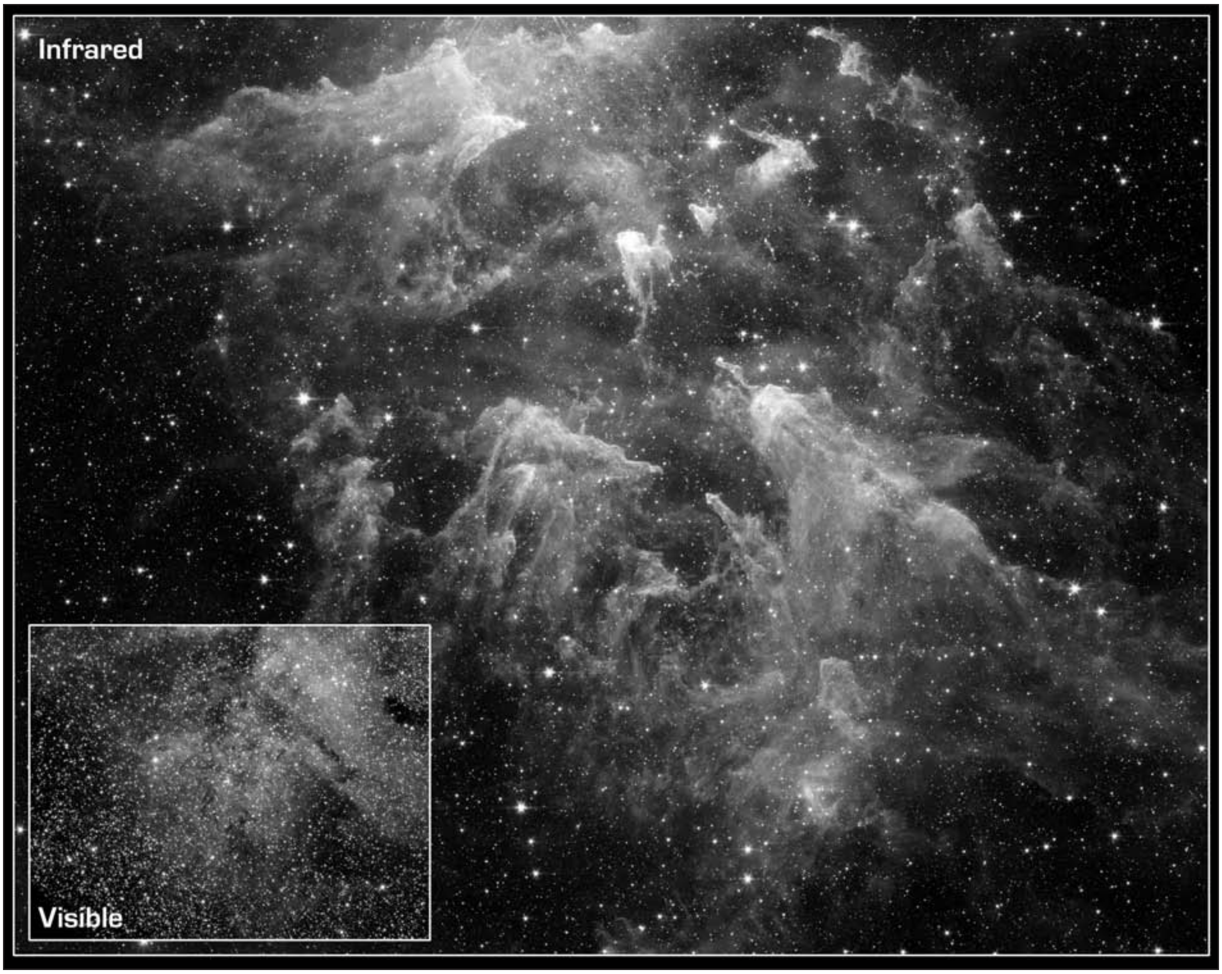
APPLICATIONS

Infrared radiation can give astronomers valuable information about the formation of stars. Stars are believed to be formed from large clouds of rotating dust and gas that condense under their own gravity. Energy released in the collapse causes the forming star, or "protostar," to increase in temperature until nuclear reactions begin. It is not until the star "turns on" in this way that it begins to emit radiation in visible wavelengths. As a result, the process of star formation is difficult to study optically. As the star begins to shine, newly created energy warms the surrounding dust, which radiates the energy away as infrared waves. The process is still not understood completely, and astronomers have learned much by the study of infrared and radio wavelengths.

Many infrared sources are clouds of dust heated by a nearby star. Infrared stars generally are either very young or very old stars, those that are associated with dust clouds. One of the early infrared discoveries was of a giant cloud of gas and dust in the constellation of Orion: the Kleinmann-Low nebula, named for its discoverers. It was found to have a mass greater than two hundred times that of the Sun, yet it is invisible at optical wavelengths. In the infrared, it outshines the Sun more than 100,000 times. It was determined to be a relatively close area of active star formation (within 1,600 light-years). Detailed studies of the Kleinmann-Low nebula in the infrared and radio bands indicate that it contains a number of young stars and clouds of dust and gas that may be in the process of collapsing to form new stars. By studying this nebula, astronomers are learning more about the process of star formation. ISO detected such "stellar nurseries" in the Milky Way and other galaxies, where, its data suggest, star formation occurs at a higher rate than astronomers expected.

One of the most exciting discoveries made by IRAS was a disk of dust grains around the star Vega. Scientists believe this disk of material may be remnants of the dust cloud from which the star formed. The theory in 1990 of planet formation suggests that a similar but smaller disk of material around the Sun provided the raw material from which Earth and other planets were formed. If the disk of material around Vega follows the same pattern, it could eventually form asteroid or planet-sized bodies. The findings from IRAS suggest that such material is common around other stars as well.

A year after the disk of material was found around Vega, a small companion object was found orbiting a faint star. The object was between thirty and eighty times the mass of Jupiter. It was too small to sustain nuclear reactions, as a star would, and some astronomers suggested initially that this object heralded the discovery of the first planet outside our solar system. It was theorized that the object was a



This infrared image of the Carina nebula, taken by the Spitzer Space Telescope, when compared with the same view in the optical range (inset) shows a stark difference in the phenomena that can be detected. The infrared image shows dust pillars where stars are forming at the "tips." These features are obscured in the optical image because of interstellar dust that absorbs visible light. (NASA/JPL-Caltech/N. Smith, University of Colorado at Boulder)

brown dwarf, an object between a star and a planet. It was an important discovery, as astronomers were finding there may be many more brown dwarf-type objects than expected. It was thought that, if these objects outnumbered visible stars, astronomers' theories regarding the amount of matter in the universe (and its eventual fate) would be in need of revision. In due course, however, it was realized that brown dwarfs could not account for all the missing mass in the universe, even as more brown dwarfs were being found by the Spitzer Space Telescope.

IRAS examined many peculiar galaxies, one of which is a galaxy known as Arp 220. IRAS found that the galaxy was emitting eighty times more energy in the infrared than in all other wavelengths. Although the object is not excessively bright at optical wavelengths, its infrared brightness would make it about as energetic as some quasars. (Quasars are extremely powerful, bright sources of energy located in a very small area at the center of a galaxy that outshine the entire galaxy around them.) Arp 220 is actually two galaxies that are colliding. While the individual stars of the galaxies are not likely to collide, huge clouds of dust and gas would collide, generating shock waves and heat by compression. This energy would be radiated in the infrared.

Researchers using IRAS employed a very rigorous observational screening process to weed out any stray infrared detections caused by charged particles. They screened out all but the sources that remained stationary over time and were repeatable. This method of observation lent itself to the discovery of some fast-moving objects that were eliminated because they moved too quickly from one observation to the next. In studying the rejected observations, scientists discovered a comet in 1983—the IRAS-Araki-Alcock—named for the satellite and for G. Iraki and E. Alcock, the independent discoverers. The comet passed closer to the Sun than any other comet in the last two hundred years, and IRAS was able to study it in detail, along with other ground-based observations. In total, six comets were discovered by the satellite, and five other known comets were studied.

Among the most important discoveries made by scientists using ISO was the signature of water around planets within the solar system and in regions of stellar formation. With regard to the latter, such a signature had been hidden from detection by the presence of dust within such forming star systems. With regard to the former, ISO determined that as much as 10 kilograms per second of water "rain" down in the upper atmosphere of Jupiter, Saturn, and Uranus. ISO's data did not answer the question as to the origin of that water. Water was also found in the thick atmosphere of Titan, one of Saturn's satellites—a tantalizing result that Cassini-Huygens scientists eagerly hoped to verify and investigate further when that probe entered the Saturn system beginning in July, 2004. Cassini found evidence of water clouds in Saturn's lowest cloud deck at a distance about 130 kilometers under the tropopause. At that atmospheric level, the local temperature is near the freezing point of water. As for the water in the gas giant planets' upper atmosphere, one leading hypothesis for the source was influx of small cometary nuclei.

ISO was also used for extragalactic studies. For the first time evidence of dust was confirmed for the otherwise rather empty space between galaxies. One particularly outstanding extragalactic finding from ISO was that intergalactic dust in a large group of more than five hundred galaxies clustered together within the constellation Coma Berenices was heavily concentrated toward the cluster's center. Determining intergalactic space is laden with very low-density dust concentrations held implications for cosmological models.

NICMOS provided a means for HST to observe celestial objects in infrared, which could be then contrasted with images of the same sources taken in visible light. One major comparison involved deep-sky observations of distant dim galaxies. The visible light images and infrared images of the same tiny areas of the sky were combined to indicate the differences between different classes of galaxies. The false-color scale used in making these survey images indicated galaxies with strong infrared emissions as reddish, while galaxies glowing more strongly in visible light appeared bluish. These sorts of survey images included blue dwarf galaxies, red elliptical galaxies, and spiral disk galaxies. Examination of those images provided insights into the populations of dust-obscured galaxies at the earliest times in galaxy formation, shortly after the big bang.

Closer to home, NICMOS was used to examine Uranus's cloud features. In visible light, Uranus reveals little of its atmospheric structure. Using infrared, however, HST's NICMOS instrument found as many as twenty clouds near a bright band in the planet's atmosphere. Wind speeds in the region were determined to be between 300 and 500 kilometers per hour.

Naturally, infrared astronomers expected great performance from the Spitzer Space Telescope, but they were pleased beyond those preliminary expectations when initial images taken by Spitzer of the dust disk surrounding a forming star, the glow of a stellar nursery, and the swirling dust in a large galaxy revealed tremendous detail. The spectroscopic capability of Spitzer revealed the signature of organic material. SST was the first observatory to detect light directly from extrasolar hot Jupiters (specifically HD 209458b and TrES-1). Views of our neighboring galaxy, the Andromeda galaxy (or

M31), taken by Spitzer clearly show the spiral arms by noting the dust lanes in them. SST studies were the first to examine the core of the Milky Way in such a way as to determine that the galaxy core has a barred structure.

Spitzer was used to determine the atmospheric temperature of the extrasolar planet HD 189733b, the first time such a measurement was made. Astronomers used Spitzer to perform surveys with long time exposures. In early summer 2008, at the American Astronomical Society meeting held in St. Louis, infrared astronomers presented an infrared image of the Milky Way that consisted of a collection of 800,000 individual images. This composite image showed the distribution of dust within the galaxy in greater detail than had previously been possible.

In early 2008 the Spitzer Space Telescope's infrared spectrometer recorded the first evidence proving the existence of water in protoplanetary disks. Observations were made of DR Tau and AS 205A, which are 457 and 391 light-years distant from Earth, respectively. Water is an essential ingredient in an evolving solar system for the possibility of an Earth-like planet forming, one that might permit the development of life. Water is also, however, important in a protoplanetary system for the formation of icy satellites around large planets in the outer fringes of the system. Water closer to the young star at the center of the protoplanetary disk could be in gaseous or perhaps even liquid form. Both of the aforementioned protoplanetary systems produced large numbers of water emission lines.

Spitzer provided insight into numerous astrophysical phenomena, one of the most bizarre being magnetars. These are stars with magnetic fields approximately fourteen orders of magnitude more intense than that of typical stars on the main sequence. Magnetars are the highly compressed remains of massive stars that went supernova, but they are hardly dead stars. In addition to their intense magnetic fields, magnetars pulsate in the X-ray portion of the electromagnetic spectrum. In late May, 2008, the Spitzer Science Center, run by the California Institute of Technology, reported on observations of the magnetar SGR 1900+14. In addition to having the usual attributes of a magnetar, this object was surrounded by a ring of material 7 light-years across that was energized by the energetic X-ray pulsations. Heated dust in the ring resulted in the ring glowing strongly in the infrared.

Although results from IRAS, HST, ISO, and Spitzer have been spectacular, ground-based telescopes are still useful for observing many infrared phenomena. Infrared observations from NASA's Infrared Telescope Facility on Mauna Kea have revealed volcanic eruptions on Jupiter's moon Io. A volcano that had been erupting at the time of the Voyager flybys in 1979 was found to be erupting still, and a new volcano was detected. Observations such as these help to gather valuable information over time that can elaborate on the findings of other missions.

CONTEXT

Infrared astronomy is part of the revolution that has been called the "new astronomy." Instruments of modern astronomers give them access to information from the entire range of the electromagnetic spectrum. This revolution has occurred mostly since the early 1960's, when it became possible to place remote detectors above Earth's atmosphere. Before then, astronomers relied for the most part on the optical range of wavelengths for their information about the universe.

It was not until 1800 that the first sign of another way to look at the universe was discovered. While analyzing sunlight by separating the white light into a spectrum, English astronomer Sir William Herschel noticed that a thermometer placed in the dark area just outside the red limit of the spectrum registered an increase in temperature. In 1881, American astronomer Samuel Pierpont Langley developed the bolometer, an electrical detector that measures heat over a broad range of wavelengths. In measuring the Sun's energy from a high altitude, Langley found that the radiant energy of the Sun extended far past the visible portion of the spectrum and far past the region that Herschel had discovered previously. Herschel had discovered the near-infrared, whereas Langley was detecting the longer-wavelength middle-infrared band.

Infrared radiation from the Sun was fairly simple to detect, but more sensitive instruments had to be developed before it was possible to detect the infrared from far distant sources. In 1856, near-infrared radiation was detected from the Moon, but it was not until the 1920's that it began to be detected from the other planets and bright stars. Available instruments were still unable to see into the far infrared. While working on superconductivity experiments in the late 1950's, physicist Frank Low began the development of more sensitive instruments. By the early 1970's, Low was among the first to attempt observations of the far infrared by leading observations aboard high-flying jets.

An infrared satellite was first proposed in the mid-1970's. NASA was facing troubled times with budget cuts, inflation, and cost overruns in other projects. It might have scrapped the project entirely except for the interest of the Dutch space agency. The Dutch had completed several successful satellite programs and were interested in collaborating on an infrared satellite. England then joined the project, which came to be known as the Infrared Astronomical Satellite program. The project was a difficult one, but the diplomatic aspects of an international collaboration helped to give the program stability, and the satellite was launched successfully in 1983. A succession of increasingly sophisticated space-based observatories expanded upon the groundwork laid by IRAS. Hubble, ISO, Spitzer, and eventually the James E. Webb Space Telescope and turned infrared astronomy into an integral component of astrophysical investigations, leading to insights into the interstellar medium, thermal processes on planets in the solar system, protoplanetary disks, nebulae, galaxy formation, brown dwarfs, and many other phenomena.

Divonna Ogier

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. This traditional treatise on introductory astronomy includes up-to-date spacecraft information such as that from NASA's Great Observatories program and rovers on Mars. Includes images that span the electromagnetic spectrum.

Dinwiddie, Robert, et al. *Universe*. New York: DK Adult, 2005. A remarkable collection of articles written by science writers and professional astronomers on a wide range of topics that span the discipline of astronomy. Heavily illustrated and filled with high-quality photographs. For the general reader.

Gregory, Stephen A. *Introductory Astronomy and Astrophysics*. 4th ed. San Francisco: Brooks/Cole, 1997. Suitable as a textbook for an introductory college-level or advanced high school course in general astronomy. Covers all topics from solar system bodies to cosmology. Some errors and issues with mathematical presentations.

Henbest, Nigel, and Michael Marten. *The New Astronomy*. New York: Cambridge University Press, 1983. Compares optical, infrared, ultraviolet, radio, and X-ray observations of well-known astronomical objects. For general readers.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find the standard treatise for introductory astronomy classes too low-level. Suitable for an audience with varied science and mathematical backgrounds. Covers all topics from solar-system objects to cosmology.

Kwok, Sun. *Physics and Chemistry of the Interstellar Medium*. New York: University Science Books, 2006. Although this text emphasizes physical processes, it is suitable for undergraduate courses in advanced astronomy covering infrared astronomy. Also addresses astrochemistry and mathematical theory on the interstellar medium.

Seeds, Michael A. *Horizons: Exploring the Universe*. New York: Brooks/Cole, 2007. A general astronomy text that also asks big-picture questions. Examines humanity's place in the universe, including physical and biological evolution.

Spitzer, Lyman, Jr. *Physical Processes in the Interstellar Medium*. New York: Wiley, 1998. Written by the astronomer after whom the

Spitzer Space Telescope is named, this work covers the physics and chemistry of the interstellar medium by frequently referencing infrared observational data. Accessible to the general reader as well as the astronomy enthusiast and student.

Verschuur, Gerrit L. *The Invisible Universe: The Story of Radio Astronomy*. New York: Springer Praxis, 2006. Provides a history of developments in radio astronomy and along the way describes the discovery of pulsars, quasars, and radio galaxies. Suitable for a general college science course as well as for astronomy majors as background information.

See also: Archaeoastronomy; Coordinate Systems; Earth System Science; Gravity Measurement; Hertzsprung-Russell Diagram; Neutrino Astronomy; Optical Astronomy; Protostars; Radio Astronomy; Telescopes: Ground-Based; Telescopes: Space-Based; Ultraviolet Astronomy; X-Ray and Gamma-Ray Astronomy.

Interplanetary Environment

Category: The Solar System as a Whole

Far from empty, the vast spaces between the Sun and its planets and out to the edge of interplanetary space constitute a dynamic environment suffused with fields of forces, crossed by swiftly moving particles, littered with debris, and penetrated by cosmic rays from outside the solar system. These phenomena endanger human technology, both in space and on Earth, but they also tell scientists much about the Sun and the origin of the solar system.

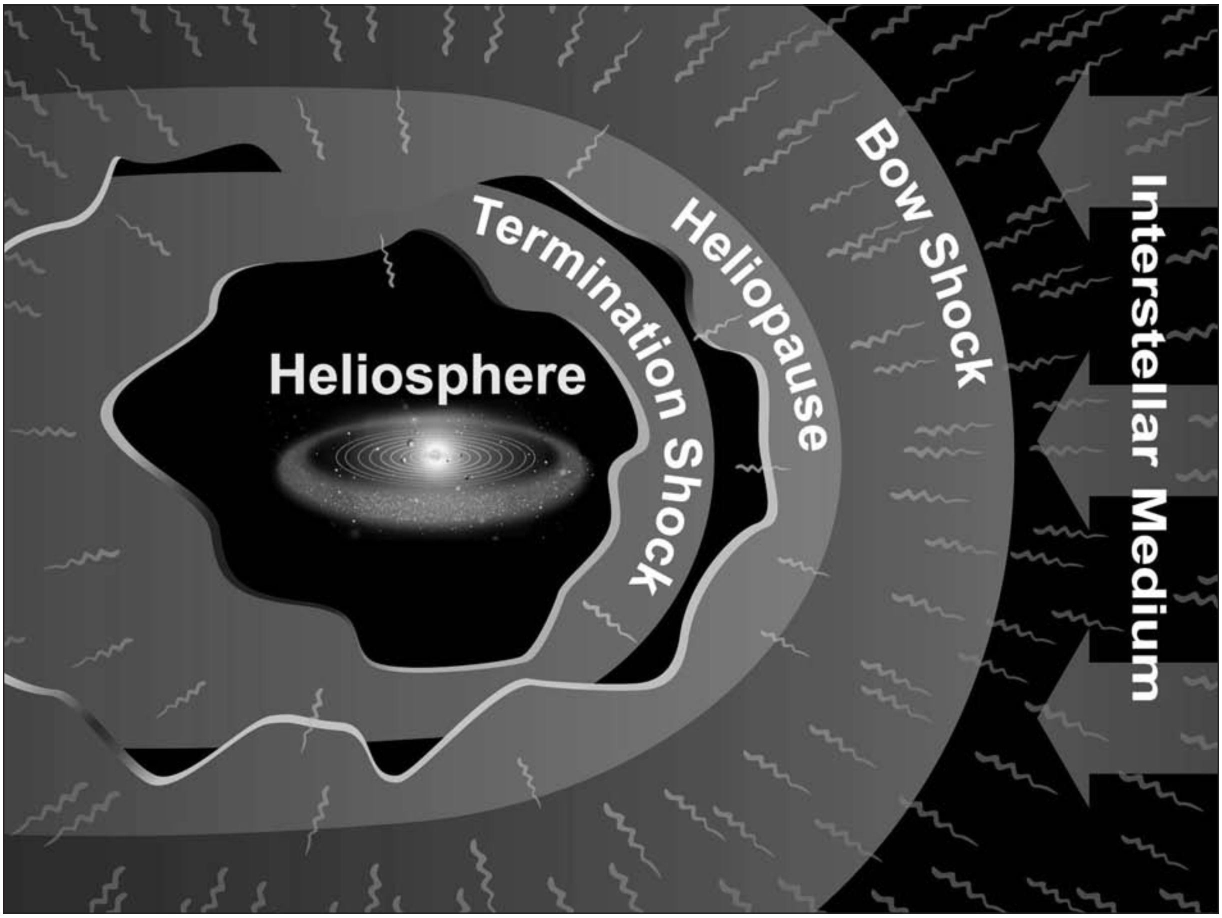
OVERVIEW

The interplanetary environment principally contains materials ejected from the Sun, and debris. Space debris comprises particles of great variety in size and composition. Anything smaller than 0.01 millimeter, astronomers call micrometeoroids or interplanetary dust. Anything larger is a meteoroid, of which many are tens of meters in size. By far the largest proportion of the debris originates when asteroids collide in the asteroid belt between the orbits of Mars and Jupiter, or from comets that swing into the inner solar system and leave behind a trail of particles eroded from them by the solar wind. Accordingly, most dust lies in the plane of the ecliptic (the disk-shaped region of the Sun and planets) in two bands: in the inner solar system, out to about 3 astronomical units (AU, the average distance from the Sun to Earth), and in the Kuiper Belt, about 10 to 40 AU. Additionally, a small amount of matter infiltrates from interstellar space as the solar system drifts through galactic gas clouds or from volcanoes on satellites, such as Jupiter's Io. Found both as chondrites (clusters of particles) and as solid chunks, interplanetary dust, like meteoroids, is rich in carbon, iron, sulfur, nickel, and silicates, but many other elements in mineral combinations have been found, including tiny spheroids of glass embedded with metal sulfides (known as GEMS). In a clear, dark sky of late evening or early morning, interplanetary dust is visible near the horizon in the direction of the Sun as a faint glow of reflected sunlight, called zodiacal light.

Although the interplanetary debris is largely gathered into two clouds, it is not static. Depending on size and location, particles move at different speeds and in different directions. The pressure exerted by solar photons pushes micrometeoroids between 0.5 and 1 micron slowly outward and eventually out of the solar system altogether. Particles larger than 1 micron are affected by radiation pressure that creates drag, which causes them to decelerate so that they fall inward toward the Sun until they are vaporized. Asteroid collisions and comets replenish the clouds. It is also the case that debris left behind by comets retains the parent body's velocity and orbit as a meteoroid stream. When the Earth passes through one of these streams, the vastly increased numbers of particles entering the atmosphere can create spectacular displays of streaking light in meteor showers, notably the Leonid showers each November and the Perseid showers each August (named after the constellations Leo and Perseus, from which they appear to arrive).

In addition to electromagnetic radiation, such as the photons of visible light, the Sun generates the gravitational field that holds together the solar system and magnetic fields. Magnetic fields accelerate a steady stream of plasma (electrons, protons, and ionized atoms) in all directions. That stream of particles is the solar wind. Near Earth the solar wind has an average density of about ten particles per cubic centimeter, a velocity of about 400 kilometers per second, and a temperature of about 100,000 kelvins, although these properties are highly variable. Particle velocity remains fairly constant beyond Earth's orbit, but density decreases as the solar wind spreads outward.

The solar wind defines the extent of the Sun's influence within its galactic neighborhood. As the plasma attenuates in the outer solar system, it eventually is slowed by the pressure of gas between stars and is deflected. The boundary where this occurs is called the termination shock. It fluctuates with the intensity of the solar wind and density of interstellar gas at between 90 and 100 AU. In December, 2004, the Voyager 1 spacecraft confirmed its existence, passing through it at about 94 AU. Voyager 1 then passed into a region of turbulence where the solar wind mixes with interstellar particles, called the heliosheath. Its outermost edge, the heliopause, marks the limit of the heliosphere,



An artist's diagram of the heliosphere, created by the solar wind, which suffuses the interplanetary environment and is slowed at the point called the termination shock, where it meets the incoming interstellar medium. This region, along with the heliopause and external bow shock, define the outer boundary of our solar system and may fluctuate somewhat with the competing pressures of solar wind and incoming radiation of the interstellar medium. (NASA) the total area the solar system. The heliosphere is something of a misnomer, however, because it assumes a teardrop shape as the solar system moves through interstellar gas clouds; its forward edge is thought to extend to between 115 and 150 AU.

Occasionally the Sun hurls outward immense bubbles of matter called coronal mass ejections (CMEs). Also known as solar storms, they typically involve between one hundred trillion and one quadrillion grams from the Sun's outer layer. By the time the typical CME reaches Earth's orbit, it has a speed, on average, of about 280 kilometers per second. It also carries with it its own magnetic field. The frequency of CMEs varies during an eleven-year cycle of solar activity. During solar maximum, Earth happens to be in the way of about seventy-two CMEs per year; at solar minimum, eight CMEs. This constitutes a minority of storms that the Sun ejects in all directions; during solar maximum the average is 3.5 CMEs per day. They can vary enormously in size and speed, the slowest moving a few dozen meters per second and the fastest 2.5 kilometers per second.

Planets with magnetic fields, such as Earth, are protected from the solar wind and CMEs. Earth's magnetic field, for instance, deflects the ionized particles so that it resides at the center of a tear-shaped bubble of relative calm. The interaction between the solar wind and Earth's southern and northern magnetic poles produces auroras, eerily rippling sheets of color in the sky. Satellites, asteroids, and planets with little or no magnetic field, such as Mars, are subject to a bombardment of their surfaces, which is deadly to life.

Sudden eruptions on the Sun's surface called solar flares emit X rays that can affect the properties of a planet's upper atmosphere. The Sun also broadcasts a constant blizzard of neutrinos. Some 70 billion of them strike every square centimeter of Earth's surface every second. However, these particles, nearly without mass, seldom interact with other matter. Although the solar wind deflects low-energy cosmic rays, the high-energy variety penetrate the interplanetary environment from unknown, distant sources; additionally, like everything in space, the solar system is bathed in the cosmic background radiation, the fading glow from the universe's origin.

KNOWLEDGE GAINED

Forces and particles of the interplanetary medium provide information about the origin of the solar system, its composition, the structure and behavior of the Sun, the evolution of planets, and the nature of other planetary systems.

The pervasive streams of neutrinos, for example, confirm theoretical calculations about the conversion of hydrogen into helium during fusion, the nuclear reaction that produces the Sun's radiant power. The gusty solar wind enables scientists not only to sample the constituent elements of the radiation; it also provides clues to the behavior of magnetic fields in the solar corona. The production of high-energy X rays by solar flares reveals the extent of those fields' power, as do CMEs. The frequency of CMEs and flares characterizes the Sun's eleven-year cycle of activity.

Most of the matter left over from the formation of the Sun and planets was long ago ejected from the solar system. However, enough remains mixed in with interplanetary debris, particularly in particles shed from comets, to provide clues about the relative abundance of elements, and specific isotopes, in the presolar gas cloud. With this information, astronomers can distinguish the unique chemical makeup of the solar system. Moreover, the Earth receives a steady rain of particles from the interplanetary dust clouds. Geologists estimate about 40,000 metric tons falls to the surface yearly.

Because the solar system's present dust clouds result from collisions between asteroids and material sloughed from comets, it is reasonable to infer that dust clouds around other mature stars evolved from the same processes. Astronomers have detected such dust clouds around about one-third of stars. Planetary systems appear to be common.

CONTEXT

Knowledge acquired about the interplanetary medium is of more than scientific concern. Understanding the interplanetary medium helps protect humanity from the dangers posed by its various contents. Meteoroids and micrometeoroids streak past Earth at speeds of 11 to 70 kilometers per second, so even the smallest carries enough kinetic energy to damage space vehicles. Many satellites and the International Space Station suffered minor punctures even though the number of micrometeoroids in near-Earth orbit is low.

The solar wind is also dangerous. Astronauts and sensitive electronic equipment must be shielded from it. Still more dangerous is the sudden hurricane of particles and magnetic fields unleashed in a CME. Astronauts then must retreat to heavily protected areas in their spacecraft, and unprotected satellites may be rendered useless. Especially large CMEs can penetrate Earth's protective magnetosphere to disrupt communications and cause power outages. Because of the Earth's tilt toward the Sun and the amount of landmass in the Northern Hemisphere, power grids there are particularly at risk. A surge of X rays from a solar flare can likewise affect Earth's ionosphere and drown out radio communications. Additionally, exposure to cosmic rays increases the risk of cancer for astronauts, a fundamental problem to overcome if there are to be long voyages, as from Earth to Mars.

The combination of relative particle densities, solar flares, the solar wind, and CMEs is called space weather. As humanity grows more dependent on technology, it also becomes more vulnerable to the vagaries of space weather. Should a fierce solar storm strike unexpectedly, it could cause chaos in civilian and military telecommunications, weather prediction, and global positioning systems. It could also cause massive power outages that would bring modern life to a virtual standstill.

As a result, space agencies in the United States, the European Union, Russia, and Japan launched a variety of space probes and orbiting observatories to monitor the interplanetary medium, particularly as it is influenced by solar activity. Some of these, such as the Solar and Heliospheric Observatory (SOHO), can detect a solar storm before it reaches Earth, giving technicians time to protect or power-down sensitive equipment and giving astronauts time to seek shelter. Ground-based observatories offer similar vigilance, some of them watching especially for large meteoroids or meteoroid steams that cross Earth's orbit.

The perils awaiting in the tenuous space between planets, as well as the wealth of information and potentially exploitable resources there, offer a lesson. Humanity exists not only in the framework of civilization and the physical environment of Earth's biosphere; its survival also requires understanding of interplanetary space.

Roger Smith

FURTHER READING

Bradley, J. P. "Interplanetary Dust Particles." In *Meteorites, Comets, and Planets*, edited by Andrew M. Davis. San Diego, Calif.: Elsevier, 2005. A thorough survey of knowledge about interplanetary particles by a physicist, with helpful graphics.

Jones, Barrie W. *Discovering the Solar System*. New York: John Wiley & Sons, 1999. This moderately technical introduction to the solar system includes a chapter on its formation and small bodies, such a micrometeoroids. Graphics and other illustrations, glossary, and bibliography.

Lang, Kenneth R. *The Cambridge Guide to the Solar System*. New York: Cambridge University Press, 2003. Many color and black-and-white photographs accompany the information-rich but generally nontechnical text in this introductory book, which includes information on the solar wind and interplanetary dust. With bibliography and sidebars devoted to special topics.

_______. *Sun, Earth, and Sky*. 2d ed. New York: Springer, 2006. Provides a thorough introduction to the Sun-Earth relationship, without recourse to mathematics, including sections on solar particles and magnetic fields. Includes a wealth of color illustrations and graphics, sidebars on special topics, a generous glossary, and a bibliography.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. New York: Cambridge University Press, 2004. Written on the level of an introductory text suitable for high school science students, this collection has a section on minor bodies of the solar system that discusses interplanetary particles and other sections on solar phenomena. Color illustrations and graphics, glossary, sidebars on special topics, and bibliography.

Peuker-Ehrenbrink, Bernhard, and Birger Schmitz, eds. *Accretion of Extraterrestrial Matter Throughout Earth's History*. New York: Kluwer Academic/Plenum, 2001. A collection of professional articles by a variety of specialists. Especially valuable, and accessible to lay readers is "The Origin and Properties of Dust Impacting Earth," by Donald E. Brownlee.

See also: Earth-Moon Relations; Earth-Sun Relations; Earth's Origin; Eclipses; Kuiper Belt; Oort Cloud; Solar System: Element Distribution; Solar System: Origins.

Interstellar Clouds and the Interstellar Medium

Category: The Cosmological Context

Although the space between stars appears quite empty by terrestrial standards, careful observation yields evidence for the existence of interstellar dust grains and more than one hundred different kinds of molecules present under a variety of conditions. Study of the interstellar medium provides important information on the life cycles of stars and evolution of the cosmos.

OVERVIEW

From the terrestrial standpoint, "outer space" begins where Earth's atmosphere ends. One thinks of the solar system as a group of eight planets together with a collection of dwarf planets, smaller bodies, moons, asteroids, and comets, which orbit the Sun. Similarly, the Sun and several hundred billion other stars orbit the center of our galaxy, the Milky Way. The fact that starlight can be seen from stars in this galaxy hundreds or thousands of light-years away would suggest that the space between the stars—the interstellar medium—is essentially empty. Careful observation has shown, however, that a variety of atoms and molecules can be found almost anywhere in the galaxy. Furthermore, the density and temperature of the interstellar medium vary greatly from place to place; in many regions, this medium forms "interstellar clouds," that may emit or scatter light or obscure the view of the stars behind them.

By terrestrial standards, interstellar regions are very nearly empty, seemingly a vacuum. Even the best vacuums attainable in the laboratory, however—about one one-hundred-trillionth of an atmosphere—contain about 100,000 molecules per cubic centimeter. This is perhaps the upper limit of density for interstellar matter. The space between stars in a galaxy is so immense, compared to the space actually occupied by the stars, that 10 percent or more of a galaxy's mass can be contained in the interstellar medium. The important characteristics of the interstellar medium are density, or number of particles per cubic centimeter; chemical composition; and temperature, which provides a measure of the speed of the molecules in the medium. The interstellar medium is said to be cloudlike in character, in that different regions have different densities. The least dense material, called the intercloud gas, has from one to ten particles per 100 cubic centimeters and an absolute temperature of about 10,000 kelvins.

The known components of the interstellar medium have been identified either through the electromagnetic radiation they emit or by their absorbing effects on the electromagnetic radiation from stars that pass through them. While Earth-based studies of the interstellar medium have had to rely on the microwave region of the electromagnetic spectrum—observed by radio telescopes—and on the visible region, satellite and rocket or balloon-based observations in the infrared, ultraviolet, and X-ray regions have contributed much to the understanding of the interstellar medium.

A few basic physical principles provide the basis for the interpretation of the electromagnetic radiation from the interstellar medium. The atoms or molecules of a gas at low density will emit or absorb electromagnetic radiation only at those energies (proportional to the frequency and inversely proportional to the wavelength) that correspond to the difference in energy between allowed quantum mechanical states. If a body of gas lies in front of a star, its chemical composition can be determined from the absorption spectrum, or pattern of dark lines seen against the continuous background spectrum of starlight when viewed through a spectroscope. The temperature of a gas cloud can be determined from the frequency at which the greatest amount of emission occurs. If the gas is in motion toward Earth, the emitted radiation is shifted to slightly higher frequencies and shorter wavelengths. If the gas is moving away from Earth, the emission is shifted to slightly lower frequencies and longer wavelengths. This Doppler shift makes it possible to determine the speed of an interstellar cloud relative to Earth.

A now famous 1995 image from the Hubble Space Telescope that redefined the popular image of the interstellar medium: Large clouds of cool interstellar hydrogen gas and dust, part of the Eagle Nebula, act as stellar incubators. (NASA/ESA/STScI/J. Hester and P. Scowen, Arizona State University)

By far, the most common chemical elements found in interstellar space are hydrogen and helium. Astronomers distinguish between H I regions, in which the hydrogen exists primarily in the form of isolated atoms, and H II regions, in which the hydrogen exists primarily in ionized form, as separated protons and electrons. H II regions can be quite warm, with temperatures of 10,000 kelvins or more, and are characterized by a red emission produced by excited hydrogen atoms. Such regions include emission nebulae such as the Trifid nebula, which is one of the most beautiful objects in the sky. H I regions are somewhat cooler than H II regions and are characterized by the emission of radiation at 21-centimeter wavelengths, which can be detected by radio telescopes.

At somewhat lower temperatures, one finds a variety of cloud types in which much of the material exists as molecules rather than as separated atoms. Diffuse clouds have a temperature of about 100 kelvins and a density of about one hundred particles per cubic centimeter; they contain a mixture of hydrogen atoms and hydrogen molecules, along with some partially ionized carbon atoms, a few types of neutral atoms, and some small molecules, including carbon monoxide (CO) and formaldehyde (H_2CO_3). Dark clouds have a temperature of about 10 kelvins and a density of ten thousand particles per cubic centimeter, with most of the hydrogen appearing in molecular form. A wide variety of molecules is found in dark clouds. So-called giant molecular clouds exist at comparable temperatures and can be up to 400 light-years in diameter.

The number of molecular species that have been identified in molecular clouds is quite large. Compounds of hydrogen, carbon, nitrogen, and oxygen are most common, with sulfur and silicon appearing in a few compounds. Molecules identified include those such as ethanol (C_2H_5OH), which are stable and even common under terrestrial conditions, as well as those such as the ethynyl radical, C_2H, which are too reactive to be isolated in the laboratory but can exist for a significant length of time at the very low densities present in molecular clouds. The largest molecule identified before 1990 is the thirteen-atom linear molecule cyanopentaacetylene ($HC_{11}N$). Since then there has been an explosion of the list. Every year it seems that more new molecules are identified than during the past year.

One region of the galaxy that is particularly rich in interstellar molecules is Sagittarius B2, near the galactic center. Nearly half of the molecules identified in the interstellar medium also are found in Sagittarius B2. Some of the most complex molecules ever detected include trans-ethyl methyl ether ($CH_3OC_2H_5$), propylene (CH_3CHCH_2), and cyanodecapentayne ($HC_{10}CN$). Some molecules are even observed in deuterated form, where ordinary hydrogen (in which the nucleus consists of a single proton) is replaced by its isotope deuterium (in which the nucleus consists of one proton and one neutron); these include heavy water (D_2O), three forms of deuterated ammonia (NH_2D, NHD_2, ND_3), and two forms of deuterated formaldehyde ($HDCO$, D_2CO).

About 1 percent of the total mass of the interstellar medium exists in the form of microscopic solid particles, generally called "dust" by astronomers. Dust grains are believed to be about one-millionth of a meter or less in size, and are irregularly shaped. The presence of interstellar dust is indicated by the absorption and scattering of starlight. When a dust cloud is illuminated from the front by a nearby star, one can see the bluish scattered light. Such an object is called a reflection nebula, the best-known example of which is associated with the star group known as the Pleiades. Scattering properties of the interstellar dust grains allow astronomers to make an estimate of their size and shape. Dust particles are more efficient in scattering blue light than red light; thus, distant stars appear somewhat redder than their actual color. Although dust grains have been recovered from the interplanetary space in the solar system, it is not certain that they are representative of interstellar grains. The relatively low concentrations of certain elements in the interstellar gas, as compared to the concentrations in stellar atmospheres, indicate that these elements—magnesium, iron, and silicon—may be prevalent in the interstellar grains. One remarkable effect of the interstellar dust is the polarization of starlight that passes through it; the interstellar medium has somewhat the same effect on starlight as a polarized sunglass lens has on light. In the case of the lens, the polarizing effect is the result of long, thin light-absorbing molecules, which are held parallel to one another in the polarizing film. The dust grains must therefore be elongated in shape, and it is probable that they are held in a parallel orientation by an interstellar magnetic field, which is relatively constant in direction over large areas of space.

The amount of matter to be found in the interstellar medium varies from one galaxy to another. Elliptical galaxies appear to contain substantially less interstellar matter than spiral galaxies, such as the Milky Way. Irregular galaxies can contain very large amounts. The Magellanic Clouds, satellites of the Milky Way visible in the Southern Hemisphere, are nearly 40 percent interstellar matter. Observations of clusters of galaxies indicate the presence of an intergalactic medium with a density of less than one particle per thousand cubic centimeters and at temperatures of more than 10 million kelvins. Some heavy elements, including iron, have been identified in this medium.

APPLICATIONS

The principal reason for astronomers' interest in the interstellar medium is that it provides insights into the formation and subsequent history of stars and hence the solar system as well. According to the big bang theory, the explosive event that marked the origin of the universe produced hydrogen and some helium, but only traces of any of the heavier elements. Almost all the carbon, nitrogen, oxygen, and other elements found in interstellar molecules, the planets, and Earth's atmosphere were produced by

The spiral galaxy Messier 81, imaged here by the Spitzer Space Telescope in the infrared, displays old stars, superheated interstellar clouds and dust, and the regions of star formation. (NASA/JPL-Caltech/S. Willner, Harvard-Smithsonian Center for Astrophysics)

nuclear reactions that took place in the interiors of stars; these molecules formed from the interstellar medium and were later returned to it. Theories of the structure of the galaxy and the life cycles of the stars must account for the present composition of interstellar matter as a result of nucleosynthesis occurring in the stars and the exchange of matter between stars and the interstellar medium.

Much of the interstellar medium exists in the form of very low-density gas at a very high temperature. A density fluctuation, perhaps induced by the "shock wave" from an interstellar explosion, can trigger a collapse of a portion of this medium into a smaller region. At first, a cooling of the material occurs as the energy released by collisions of particles is radiated into space. Collisions of atoms with one another and with the dust grains result in the formation of molecules. In the Milky Way, this process appears to occur primarily in the spiral arms, which is also the location of the most recently formed stars. When cloud density becomes sufficiently great, the cloud becomes opaque to the passage of electromagnetic radiation, causing the released energy to become "trapped"; the cloud begins to warm and eventually forms a protostar, usually with an associated H II region and emission nebula. Gravitational collapse and the associated warming continue until the temperature becomes sufficiently great to allow nuclear fusion to ignite.

Nearly all nuclei in the universe, other than hydrogen and helium, were formed in the interiors of stars. The mechanism of formation is somewhat different for elements containing up to about sixty protons and neutrons (that is, up to the iron-nickel-cobalt group) and for larger nuclei, which are generally much scarcer. For the greater part of a star's lifetime, the principal nuclear reaction is the fusion of hydrogen to form helium. Once hydrogen has been substantially depleted, further gravitational collapse leads to additional warming, igniting the helium to form carbon, oxygen, and other small nuclei. If the star has sufficient mass, it may go through several additional stages of collapse and ignition, with the formation of still heavier elements. While the core of the star becomes warmer and more compact, the outer regions expand and cool so that eventually the star enters a red giant phase. The temperature in the outer extremes of a red giant are cool enough to allow the formation of molecules and possibly dust grains, which, since they are so far from the stellar core, may be able to escape the relatively weak gravitational field at the stellar surface. Planetary nebulae are H II regions that may represent a late stage in this process.

A number of other processes result in the release of matter from stars back to the interstellar medium. Stars of the Sun's mass or smaller typically have coronas, outer gaseous layers with temperatures of about a million kelvins, in which particles have sufficient speed to escape into space. For very large stars, the radiation pressure of light leaving the star is responsible for the stellar wind, a release of matter from the outer layers of a star. In binary star systems, the capture of matter from one star by the other can result in a nova, an explosion that ejects much stellar material. The sudden gravitational collapse of a large stellar core results in a supernova, which produces elements heavier than iron and returns them to the interstellar medium.

One other significant source of nucleosynthesis is the collision of existing nuclei with cosmic rays. Cosmic rays are particles, almost always protons or other nuclei, traveling at immense speeds. The interaction of cosmic rays with the interstellar medium provides a means of studying nuclear reactions occurring at very high energies, including the fragmentation of heavy nuclei, which appears to be the only source of some of the less abundant isotopes.

CONTEXT

The existence of interstellar dust and gas has played an important role in astronomers' study of the Milky Way. The belief that the broad band of light in the night sky, called the via galacta by the ancient Romans, is actually a collection of an immense number of stars of which the Sun was a member has been generally accepted since the beginning of the twentieth century. The problem was to determine the shape and size of this collection and the Sun's position in it. In the early 1900's, the Dutch astronomer Jacobus Cornelis Kapteyn conducted a survey of the distribution of stars of different magnitudes in different parts of the sky and concluded that the Sun appeared to be at the center of this distribution and that the number of stars in a given volume of space diminished with increasing distance from the Sun in any direction. This view was challenged, however, in 1917 when the American astronomer Harlow Shapley published a study of the distribution of globular clusters, large groupings of up to hundreds of thousands of stars, and showed that these clusters appeared to be centered on a point several thousand light-years from the Sun. The discrepancy between these observations was resolved in 1930 when the Swiss American astronomer Robert Julius Trumpler showed that the interstellar dust obscured the view of more distant stars and that Shapley's method, based on large collections of stars, was more reliable.

Absorption lines caused by interstellar molecules were first identified in 1904 by the German astronomer Johannes Franz Hartmann, but the first identification of an interstellar molecule, the methylidyne radical, CH, did not occur until 1939. In 1951, the American William Wilson Morgan made the first observations of H II regions in the Milky Way, identifying them with the spiral arms. At the same time, Edward Mills Purcell and Harold Ewen at Harvard University were able to detect the 21-centimeter radiation of atomic hydrogen using the techniques of radio astronomy. By measuring the Doppler shift of the 21-centimeter radiation, astronomers have been able to construct a map of the Milky Way.

Interest in the chemistry of the interstellar medium increased substantially with the discovery in 1963 of the first oxygen-containing species, the hydroxyl radical, OH, and the subsequent discoveries of interstellar water, H_2O, and ammonia, NH_3, in 1968. Later years saw the discovery of increasingly complex molecules, giving credibility to the notion that interstellar chemicals may have played some part in the origin of life on Earth or possibly elsewhere in the universe. Ever more complex molecules are being discovered primarily through radio astronomy and infrared observations.

Donald R. Franceschetti

FURTHER READING

Arny, Thomas T., and Stephen E. Schneider. *Explorations: An Introduction to Astronomy*. 5th ed. New York: McGraw-Hill, 2007. A general astronomy text for the nonspecialist. Includes an interactive CD-ROM and a companion Web site.

Dudly, W. W., and D. A. Williams. *Interstellar Chemistry*. New York: Academic Press, 1984. Provides a fascinating record of the extensive information that has been accumulated about the chemical and physical characteristics of the interstellar clouds. For more advanced readers.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory text covering the field of astronomy. Contains descriptions of astrophysical questions and their relationships.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find the standard treatise for typical astronomy 101 classes too low-level. Suitable for an audience with varied science and mathematical backgrounds. Covers all topics from solar-system objects to cosmology.

Kwok, Sun. *Physics and Chemistry of the Interstellar Medium*. New York: University Science Books, 2006. Although this text emphasizes physical processes, it is suitable for undergraduate courses in advanced astronomy. Also provides astrochemistry and mathematical theory on the interstellar medium.

Sagan, Carl. *Cosmos*. New York: Random House, 1980. This classic volume by the prominent astronomer and popular writer includes a chapter on the lives of the stars. Presents the formation of the chemical elements in a highly entertaining and memorable fashion. A companion video series of the same name is available.

Spitzer, Lyman, Jr. *Physical Processes in the Interstellar Medium*. New York: Wiley, 1998. Written by the astronomer after whom the Spitzer Space Telescope is named, this work covers the physics and chemistry of the interstellar medium by frequently referencing infrared observational data. Accessible to the general reader as well as the astronomy enthusiast and student.

Tielens, A. G. G. M. *The Physics and Chemistry of the Interstellar Medium*. New York: Cambridge University Press, 2005. A comprehensive presentation of the physics and chemistry of the interstellar medium for undergraduates through professional astronomers. Includes a detailed bibliography.

Verschuur, Gerrit L. *The Invisible Universe: The Story of Radio Astronomy*. New York: Springer Praxis, 2006. In addition to providing a history of developments in radio astronomy, this volume covers a great deal of infrared astronomy. Discusses the nature of interstellar nebulae and the detection of molecules in the interstellar medium. Suitable for a general science course in college as well as for astronomy majors as background information.

Wynn-Williams, Gareth. *The Fullness of Space: Nebulae, Stardust, and the Interstellar Medium*. Cambridge, England: Cambridge University Press, 1992. A comprehensive survey of interstellar matter and the interstellar medium, including astronomical observations in regions of the electromagnetic spectrum besides the visible. Written for laypersons, astronomy buffs, and students.

See also: Big Bang; Cosmic Rays; Cosmology; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; General Relativity; Milky Way; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Universe: Evolution; Universe: Expansion; Universe: Structure.

Io

Categories: Natural Planetary Satellites; The Jovian System

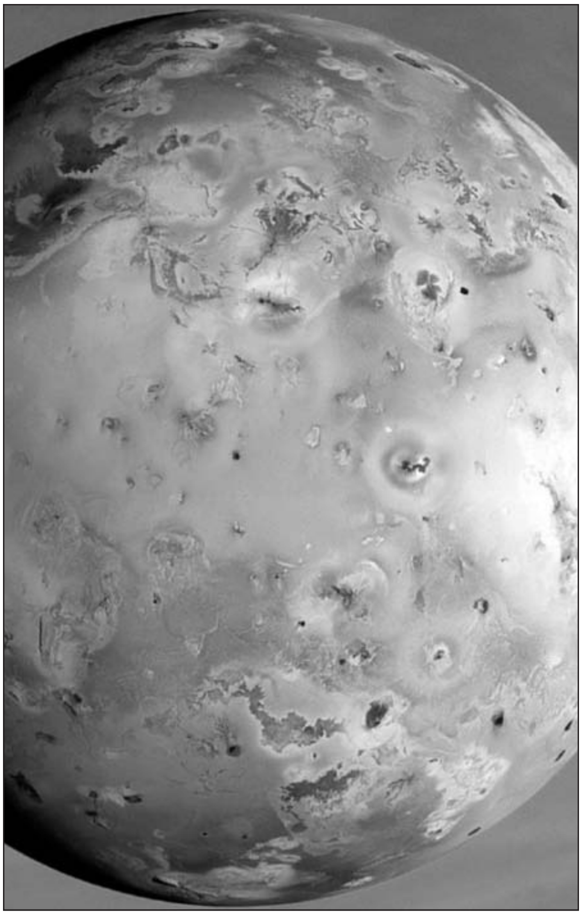
Io is the innermost of four large satellites orbiting Jupiter, the largest planet of the solar system. Io is the most volcanically active body in the solar system. Tidal friction occurs constantly on Io, heating its core. Internal thermal energy is vented through immense volcanoes that spew sulfur and sulfur components into space which fall back, resurfacing the satellite. Io has one of the youngest surfaces in the solar system.

OVERVIEW

Until 1979, Io was known as little more than a pinpoint of light, even when seen through large telescopes. Io is one of the four satellites of Jupiter that were first observed telescopically by Galileo in 1609. With roughly the same size, density, and surface gravity as Earth's Moon, Io was expected to have similar features. Earth-based spectroscopic observations during the 1970's, however, raised speculation that Io was quite different. Suspicions were dramatically confirmed by the flybys of Voyagers 1 and 2 in 1979 and by multiple encounters of the Galileo spacecraft between December, 1995, and September, 2003. The closest encounter with Io, early in the Galileo mission, was at a distance of 22,000 kilometers; the probe approached no closer because of concerns about the intense radiation environment the spacecraft would encounter. During the late stages of the extended Galileo mission, however, the spacecraft passed within a mere 180 kilometers of Io's surface; at this late point in the mission, the science return was worth the risk to the spacecraft's health.

Voyager 1 made the remarkable, and completely unsuspected, discovery that active volcanoes dot the surface of this planet-like Jovian moon. From Voyager and Galileo evidence, augmented by Earth-based observations, scientists identified Io as a world dominated by volcanic eruptions. Most of the moon's surface features are transformed daily by heated liquid, with gaseous emissions onto the surface and into the otherwise nonexistent atmosphere. On Io, impact craters are not formed as they are on the other, extremely cold, satellites of the solar system. Io's impact features are apparently absorbed by molten lava flows that extend across Io's surface. Volcanic activity is not sporadic but virtually continuous. Of the eight active volcanoes observed by Voyager 1, seven were still spewing gaseous plumes when viewed by Voyager 2 four months later. Pictures taken by the two spacecraft portrayed these huge eruptions against the backdrop of the black sky over Io's limb (its visible horizon) and from above the red-orange surface with its active calderas. However, Io's volcanic activity changed in the time between the Voyager flybys and the arrival in orbit of the Galileo spacecraft on December 7, 1995.

With its image enhanced by spacecraft instruments, Io looked like a giant pizza, with wide plains of different hues punctuated by darker and lighter active regions. The latter are



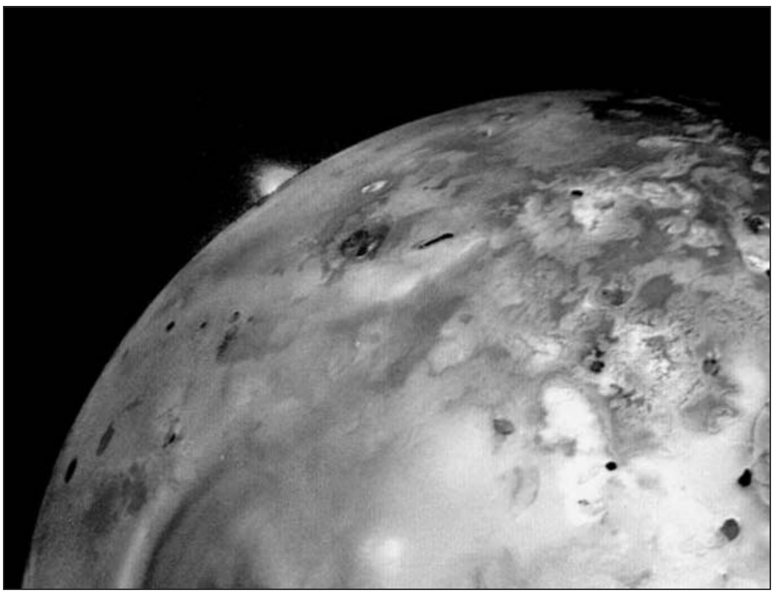
The Galileo spacecraft captured this high-resolution image of Jupiter's moon Io from about 294,000 kilometers in March, 1998. (NASA/JPL/University of Arizona)

calderas, which dot the surface, at least two hundred of them, each having a diameter of more than 20 kilometers. The largest include eleven observed plume-emitting volcanoes named by the International Astronomical Union for mythological gods of early and primitive religions. Two of these, separated along the Loki fissure, are associated with a lava lake 200 kilometers wide, known as Loki Patera, which apparently is the major outlet for the planet's internal heat. Its temperature, like those of other Io "hot spots," averages about 300 kelvins. The hot spots contrast with the remaining 98 percent of the surface, which at 130 kelvins is considerably colder, as would be expected for an atmosphere-poor body so far from the Sun.

Mountains tend to cluster near the polar regions on Io. Some have peaks as high as 10 kilometers, but they do not appear to have been formed by plate tectonics (the shifting of continental geologic structures). They lack cone-like tops and could not have been formed by recent volcanism. It is speculated that some of Io's upper crust may detach and float about the molten plains in a manner analogous to icebergs. Also, erosion scarps form near emission calderas. The fluid surface suggests that Io might receive a new surface 10 micrometers thick every year, making it unique in a solar system of much older, inactive, heavily cratered planets and satellites.

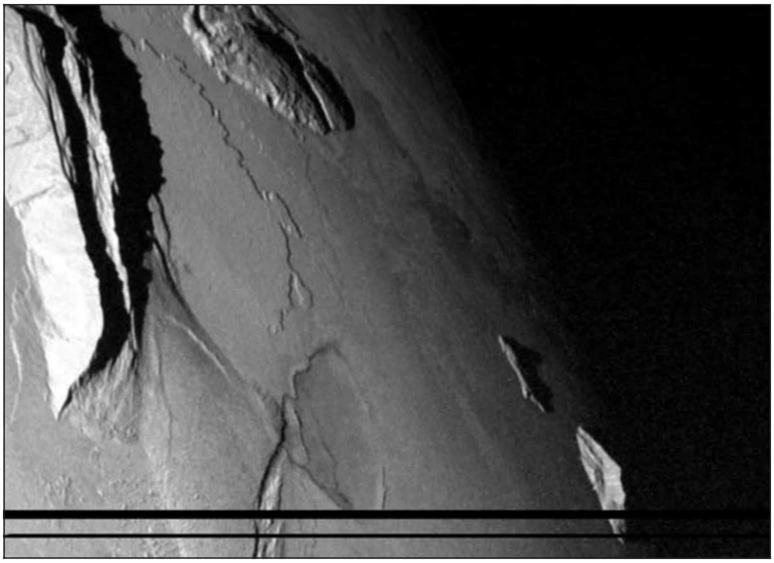
The chemistry of Io helps explain this satellite's dynamic volcanism. Previously thought to have a solid interior, like other satellites of the solar system, Io appears to have a molten silicate core. Planetary geologists hypothesize that about four billion years ago, heated sulfur dioxide lying just below the surface became the driving force for the volcanoes, ejecting Io's internal heat in gaseous eruptions similar to geysers. Long-lived eruptions, like Loki, eject materials ballistically at 0.5 to 0.6 kilometer per second. Short-duration powerhouses, typified by Pele (and possibly Aten and Surt, seen by Voyager 2) do so at twice that velocity. These eruption rates are significantly higher than that of Earth's volcanoes (0.1 kilometer per second). Sulfur compounds are ejected as majestic mushroom-shaped plumes to a maximum height of 300 kilometers, enabling the lighter compounds, such as water and carbon dioxide, to escape into space. The heavier compounds, such as sulfur and sulfur dioxide, fall back to the surface as frozen, whitish, snowlike matter. Flows of molten matter therefore are low-viscosity sulfur and sulfur compounds rather than silicate rock lavas typical of Earth's volcanoes. Io's surface color results from the various sulfur compounds.

Io's geologically active behavior is caused by its proximity to massive Jupiter and to sister Galilean satellites Europa and Ganymede. Jupiter's gravitational pull causes Io to "flex" along its axis 10 kilometers toward Jupiter, while the combined attractions of Europa and Ganymede cause torques that give Io a slightly eccentric, noncircular orbit. The result is two opposing tidal forces stretching Io from within as it orbits Jupiter every 1.77 days. The ensuing friction raises Io's internal heated power to 60 to 80 trillion watts, partially melting the silicate compounds of the crust and generating volcanic eruptions. Furthermore, because Io's orbit lies entirely within Jupiter's radiation belts, the satellite is bombarded by charged particles and affected by the powerful electrical currents produced by Jupiter's magnetic field. These phenomena also influence Io's internal heating, as does the spontaneous radioactive decay of isotopes, which is typical of all planetlike bodies.



This amazing image of Io, taken by Voyager 1 in March, 1979, shows the eruption of a huge volcano at the upper left horizon—marking the first sighting of an active volcano beyond Earth. (NASA)

Io's volcanic emissions account for its irregular atmospheric pressure, first detected by Pioneer 10 in 1973. Atmospheric pressure variations result from the heat differential associated with the anomalous hot spots and the typically cooler surface. Earth-based observations in 1974 and 1975 detected a "cloud" of neutral sodium and potassium extending along Io's orbit for more than 100,000 kilometers. This observation was explained as a "sputtering" process whereby potassium atoms are ejected into space from Io because of the impact of charged particles from Jupiter's magnetosphere striking Io's surface. Their rate of ejection is greater than 10 kilometers per second, well above the necessary escape velocity of 2.5 kilometers per second. By comparison, Io's most active volcano, Pele, has an ejection speed of only 1 kilometer per second. Volcanoes, then, are an indirect rather than a direct cause of this sodium-potassium cloud; they bring the elements to the surface in molten and gaseous states for emission into space.



In February, 2000, the Galileo spacecraft took this image of Io at a resolution of about 335 meters. To the left, Mongibello Mons rising about 7 kilometers from the surface. (NASA/JPL/University of Arizona/Arizona State University)

In 1976, additional Earth-based observations revealed a plasma "torus," or faint ring of excited glowing gas, belonging to Io's orbit. This torus occupies space within Jupiter's magnetosphere and results from the sputtering process. New electronic cameras and filters carried aboard the National Aeronautics and Space Administration's Kuiper Airborne Observatory recognized sulfur in 1981 and oxygen in 1982, as well as sodium and potassium, escaping from Io. The entire cloud assemblage supplies the torus with raw materials for further breakdown into discrete atoms and ionization. Thus energized, these ionized elements join the Jovian radiation belt that helped create them. The discovery of the torus was as unexpected as the volcanoes. Pioneer 10 and Voyager 1 flew directly through the torus in 1973 and 1979, respectively, but provided only knowledge supplementary to the major data obtained through continuous Earth-based monitoring. On its way to orbit insertion in December, 1995, the Galileo spacecraft flew a relatively safe distance from Io, one that had originally been considered to be the closest that Galileo would ever get to this innermost Jovian satellite, and then passed quickly through the plasma torus. No significant radiation damage was incurred by the spacecraft.

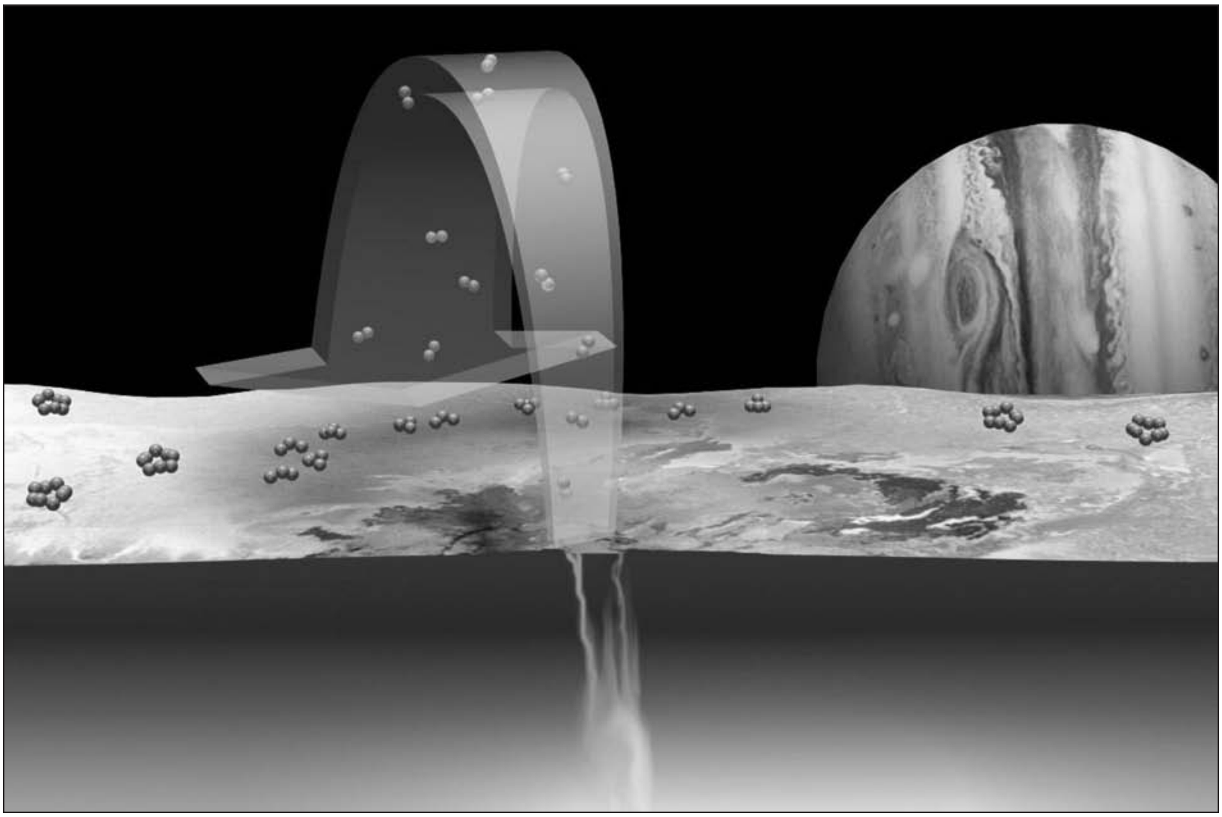
The volcanoes of Io offer the greatest promise for resolving the details of the complex relationship between Jupiter and Io. Of the eleven volcanoes discovered by Voyagers 1 and 2, the two-part vent of Loki is the most important. With a height of about 225 kilometers and a width of more than 430 kilometers, Loki and its lava lake, Loki Patera, appear to be the major outlet for Io's internal heat, as suggested by thermal emission polarization measurements in 1984. The thermal output from Io's greatest volcanoes, Pele (305 kilometers high and 1,200 kilometers wide), Surt, and Aten, is also significant. One of the three apparently ceased eruptions in 1986, according to Earth-based observations. The rest are 100 kilometers or less in height. Known calderas make up about 5 percent of Io's surface. Some, like Loki and Pele, have asymmetrical plumes and surface flows, which probably are consequences of irregular vent shapes. Others, like Prometheus, have symmetrical, fountainlike plumes and circular flows.

The Galileo spacecraft confirmed the massive scale of Io's volcanism, detecting a fresh volcanic deposit the size of Arizona and establishing that the satellite is rich in silicates. Increasingly sophisticated Earth-based instruments and orbiting telescopes have provided additional analysis. During eclipses in 1985, ice-covered Europa passed through Io's shadow, reflecting sunlight that revealed the uniform distribution of sodium about Io. The Hubble Space Telescope spotted an active plume in June, 1997, which Galileo also detected.

KNOWLEDGE GAINED

Principally from Voyager 1 and Earth-based observations, Io was revealed to be the only volcanically active satellite of Jupiter. Volcanic gases and molten lava flows were seen being emitted from eleven major fissures. The most notable is the large dual vent of Loki Patera, which appears to be the focal point of the satellite's heat emissions from its interior. Because these eruptions are continuous, the ejected heavier compounds of material steadily move across the surface, eroding low-lying scarps and erasing the craters formed by objects striking Io. At least one lava lake was discovered, along with mountains near the poles. Emissions appear to be generated by Io's molten interior, from which sulfur compounds are emitted onto the surface and into the atmosphere. The heating mechanism for this activity is apparently internal friction caused by the gravitational pull of massive Jupiter on one side, and the satellites Europa and Ganymede on the other.

Because Io lies within Jupiter's magnetosphere, charged particles strike the satellite's volcanic surface, causing jets of potassium, sodium, sulfur, and oxygen to be ejected into its atmosphere—a "sputtering" process that feeds a cloud of those neutral elements. This cloud in



An artist's conception of the role of Io's sulfur, which emerges from volcanoes and lands on the surface (see the arching arrow), where it is cooled into S_3, S_4 (both pinkish in color), and sulfur, S_8, which gives the moon its characteristic yellow hue. (NASA/JPL/Lowell Observatory) turn supplies raw materials for a torus of excited gas along Io's orbital path.

Determination of Io's bulk density and moment of inertia has revealed the satellite to be a differential body composed of a silicate mantle with a metallic core that makes up as much as half of Io's radius. Comparative studies of the four Galilean satellites indicate that they share a number of similarities in their cores as well as differences due to their environment. Io, being active, has lost any significant icy shell it might have had over its core. Ganymede and Callisto, being colder and far less active, have thick, icy shells over their cores with craters on the icy surfaces of each. Europa has activity that gives it an icy crust that breaks up and flows over what is believed to be a liquid ocean beneath that crust.

As a result of Galileo's repeated observations of Io, hundreds of volcanic sites were identified, as were about two hundred dark surfaces believed to be fresh silicate lava. The orbiter's near-infrared mapping spectrometer and solid-state imager identified a hundred active hot spots through thermal emission.

CONTEXT

Io emerged from the Voyager 1 and 2 missions as unique, not only among the satellites of Jupiter but also among all planets and satellites within the solar system. Io is recognized as the most volcanically active planetary body in a solar system where only a few other worlds display volcanism, and many of those display a type of cryovolcanic activity quite different in nature from Io's volcanoes. Io's volcanism is of particular interest because, like Earth, it is a dry body with a molten interior, and its sulfur-enriched chemistry may mimic volcanic conditions that existed during Earth's early history. Earth's active volcanoes convert water to steam for geothermal output. Io's sulfur-based volcanism provides an active laboratory for the study of planetary evolution, because volcanic eruptions constantly resurface its crust.

Io exists well within the Jovian magnetosphere. Electromagnetic fields affect Io's surface, allowing lighter elements emitted through its volcanic vents to escape into the atmosphere and feed the torus that encircles Jupiter as part of the radiation belt. The Jupiter-Io connection serves as a laboratory for the study of large-scale magnetic forces in the solar system. One of Io's effects on the Jovian system is a very gradual slowing of the rotation of Jupiter and erosion of the orbits of Europa and Ganymede.

Because active satellites like Io had not been anticipated before the 1979 Voyager flybys, their study has enhanced the evolving field of comparative planetology. Io was joined by Enceladus and Triton as satellites in the outer solar system that display unexpected volcanic activity.

Naturally, Io remains a high-priority target in planetary science for further study by robotic spacecraft. However, the radiation environment makes it difficult to dispatch probes into close proximity to the satellite. Human exploration is highly unlikely. There is a scene in Peter Hyams's film *2010: The Year We Make Contact* (1984) in which the wayward *Discovery* and its deactivated HAL 9000 computer are found adrift near Io. Spacewalking astronauts then transfer from a Russian spacecraft called the *Leonov* to enter and activate *Discovery*. In reality, the intense radiation environment would have provided such spacewalking astronauts a lethal dose well before they could return to the safety of another ship.

Clark G. Reynolds

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. Cambridge, England: Cambridge University Press, 2007. A comprehensive work about the biggest planet in the solar system. A series of articles provided by recognized experts in their fields of study. Excellent repository of photography, diagrams, and figures about the Jupiter system and the various interplanetary missions that have unveiled its secrets.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. A richly illustrated summary of early space-age discoveries that radically revised knowledge of the solar system, particularly useful in tracking initial reactions of scientists to those discoveries. Major features of Io and its volcanoes are covered.

Cole, Michael D. *Galileo Spacecraft: Mission to Jupiter*. New York: Enslow, 1999. Provides a full description of the Galileo spacecraft and science returns through the primary missions. Particularly good at describing mission objectives and goals. Suitable for a younger audience.

Fischer, Daniel. *Mission Jupiter: The Spectacular Journey of the Galileo Spacecraft*. New York: Copernicus Books, 2001. Thoroughly explains all aspects of the science and engineering of the Galileo spacecraft. Particularly good are the discussions about the nature of the Galilean satellites. Suitable for a wide range of audiences.

Geissler, Paul E. "Volcanic Activity on Io During the Galileo Era." *Annual Review of Earth and Planetary Sciences* 31 (May, 2003): 175-211. The definitive work describing the physics and planetary geology of volcanoes on Io. Provides a complete picture of Voyager and Galileo spacecraft results.

Greeley, Ronald. *Planetary Landscapes*. 2d ed. London: Allen and Unwin, 1994. A brief but instructive photographic examination of imaged planetary surfaces of the solar system, including a treatment of volcanism that sheds light on Io's unique volcanic processes.

Harland, David H. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. New York: Springer Praxis, 2000. This book provides virtually all of NASA's press releases and science updates during the first five years of the Galileo mission. Contains enormous numbers of diagrams, tables, lists, and photographs. Provides a preview of the Cassini mission. Published before the completion of the Galileo mission.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Takes a comparative planetology approach rather than including separate chapters on individual planets of the solar system.

Lopes, Rosaly M. C., and John R. Spencer. *Io After Galileo: A New View of Jupiter's Volcanic Moon*. Heidelberg: Springer, 2007. A volume in the Springer Praxis Space series, this book summarizes the knowledge gained by Galileo. Suggests new investigations needed to explain those questions that remain about the volcanism of this Jovian moon. Technical.

Morrison, David. "The Enigma Called Io." *Sky and Telescope* 69 (March, 1985): 198-205. An updated summary of original Voyager 1 and 2 data collected between 1979 and 1984, including contemporary information from Earth-based instruments, by a leading authority on the subject. Special attention is given to Loki Patera and the Jovian nebula, or thin gaseous torus, generated by Io.

Morrison, David, and Jane Samz. *Voyage to Jupiter*. NASA SP-439. Washington, D.C.: Government Printing Office, 1980. The official account of the Pioneer and Voyager flybys of the Jovian system, covering the day-to-day revelations from each mission. The most notable is the dramatic discovery of Io's volcanoes by Voyager 1. Lavishly illustrated.

See also: Callisto; Europa; Ganymede; Jovian Planets; Jupiter's Magnetic Field and Radiation Belts; Jupiter's Satellites; Neptune's Satellites; Planetary Satellites; Planetary Tectonics; Saturn's Satellites; Uranus's Satellites; Venus's Volcanoes.

J

Jovian Planets

Categories: Planets and Planetology; The Jovian System

Jupiter, Saturn, Uranus, and Neptune are called the Jovian planets. These "gas giants" have a mass 15-320 times greater than Earth, are of very low relative density, are mainly fluid (gas and liquid), and are composed of relatively light elements such as hydrogen and helium. All of them are surrounded by ring systems, a host of diverse satellites, and complex magnetospheres.

OVERVIEW

The four Jovian, or "gas giant," planets are totally different geologically and physically from the terrestrial planets, Mercury, Venus, Earth, and Mars. Massive gaseous and liquid bodies composed primarily of hydrogen and helium, the Jovian planets are relatively rapid rotators. Each rotates about its own axis in less than twenty-four hours. The atmospheres of the Jovian planets—the feature that dominates observational work done on these planets—are very similar to one another in composition. Hydrogen represents about 90 percent of the atoms present, with helium making up the bulk of the remaining atmospheric gases. Methane and ammonia are also present, although ammonia on the two colder planets Uranus and Neptune has most likely precipitated out of the atmosphere. Weather systems that dominate these atmospheres, particularly in the case of Jupiter and Saturn, consist of rapidly rotating belts and zones that are visible from Earth. In the case of Jupiter, wind speeds on the order of 300 kilometers per hour are common, while on Saturn winds of two to three times that speed have been measured. Note that on Earth, hurricane-force winds rarely exceed 150 kilometers per hour. Ironically, wind speeds in the colder Uranus and Neptune are even higher than those seen on Jupiter and Saturn.

Both Jupiter and Saturn are much hotter than might be expected in view of their distances from the Sun. They have their own heat sources deep within their planetary interiors and thus are able to produce extensive thermal cells to drive high-speed winds. The nature of the heat within these planets is not entirely evident, and there is some evidence that even in the case of Uranus there may also be a heat-driven weather system resulting from a much more modest heat source on the planet. Uranus displays little atmospheric structure in visible light, but in ultraviolet there are some features. Images from Voyager 2 have shown that Neptune has an actively driven weather system. Dark storms and white streakers are seen to evolve in short time frames.

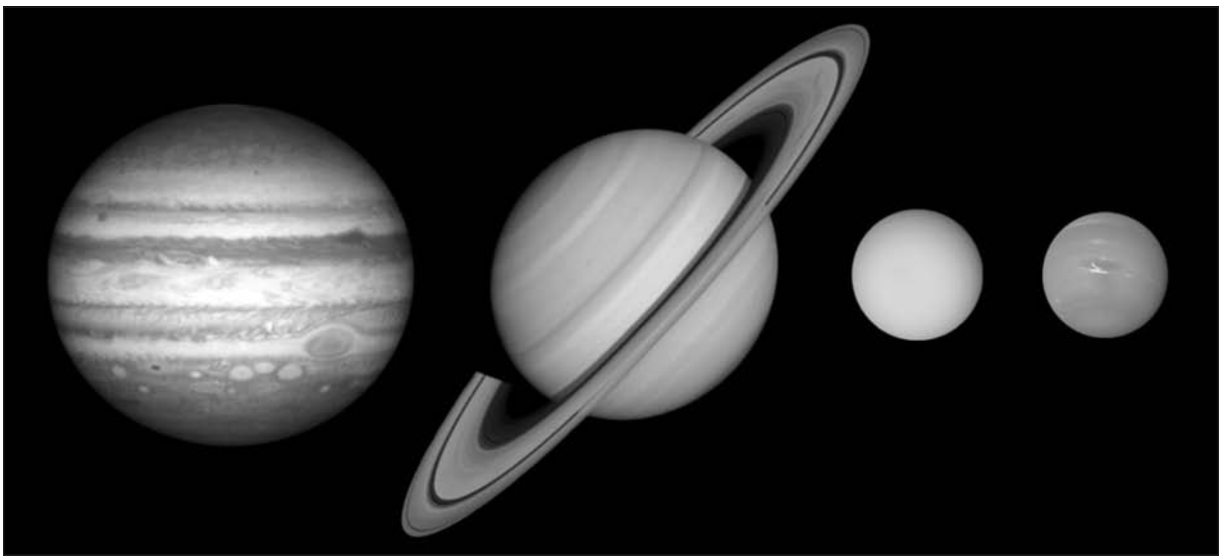
Although there are no measurements to indicate what lies below these turbulent atmospheres, there is indirect evidence that toward the center of a typical Jovian planet, pressures become higher. A portion of the interior is liquid. Toward the center of both Saturn and Jupiter, a very unusual state exists, that of liquid metallic hydrogen. This liquid metallic state would enhance both the thermal and electrical conductivity of the planetary interiors and no doubt is largely responsible for the strong magnetic fields associated with Jupiter and Saturn. Pressures necessary to create liquid metallic hydrogen are on the order of millions of times the atmospheric pressure at the surface of Earth. Although they can be re-created in tiny cells in the laboratory, no such pressures have been sustainable in large-scale systems for prolonged periods of time on Earth. Thus, the liquid metallic hydrogen layer in the interiors of Jupiter and Saturn have effects that are not yet fully described. Uranus and Neptune probably have no such layers, for their masses are not great enough to produce such enormous interior pressures.

If one were proceeding inward toward a Jovian planet's center, one would next approach the core of that planet. Theorists disagree as to what might exist there, but the dominant opinion is that the cores would be largely solid and would contain relatively heavy metals such as iron, or they may be silicon in some high-pressure phase. In the case of Jupiter, such a solid core might have a mass 20 times greater than that of Earth, but this is a small fraction of the total mass of this planet, which is 320 times that of Earth. Uranus and Neptune might have solid cores on the order of several Earth masses, while that of Saturn would be about 5 to 10 Earth masses.

Little or nothing has been learned experimentally about planetary interiors; this is true even for Earth and its fellow terrestrial planets. Modeling planetary interiors, particularly on the scale necessary in the cases of Jupiter and Saturn, requires knowledge of pressure effects on bulk matter at pressures of millions of atmospheres and at alleged temperatures of 50,000 kelvins or hotter. It is known, however, that all the Jovian planets have a very low density, that is, a low specific gravity. Specific gravity is a measure of relative density, using water as a unit of 1 gram per cubic centimeter or 1,000 kilograms per cubic meter. Saturn has a specific gravity or relative density of 0.7; this means that it would float if one could find an ocean of water big enough in which to place it. It is by far the least dense of all the planets. Jupiter has an average density of 1.3 grams per cubic centimeter, while Uranus and Neptune have average densities of 1.2 and 1.7 grams per cubic centimeter, respectively. A typical terrestrial planet has a specific gravity of about 5. Earth's density is on average 5.5 grams per cubic centimeter. Overall, such relatively low density measurements indicate the predominance in Jovian planetary structures of light elements such as hydrogen and helium.

Jupiter has a sizable magnetosphere. Its strong magnetic field is about ten times as intense as that of Earth. Jupiter's magnetosphere, which consists of trapped charged particles in amounts that would be lethal to humans, is so large that Saturn, which is 9.5 astronomical units (AU) from the Sun, passes through it. Saturn is almost twice as far from the Sun as is Jupiter (at about 5.2 AU), yet its magnetosphere is very strongly influenced by that of Jupiter. Saturn itself has a magnetic field slightly larger than that of Earth.

All the Jovian planets rotate about their own axes rapidly in relation to the terrestrial planets, which take at least twenty-four hours to make one rotation. (Earth is the fastest rotating terrestrial planet.) All Jovian planets thus exhibit some degree of oblateness. Saturn has an



A composite image of the four gas giants known as the Jovian planets (left to right): Jupiter, Saturn, Uranus, and Neptune. (Lunar and Planetary Institute)

Facts About the Jovian Planets

	Jupiter	*Saturn*	*Uranus*	*Neptune*
Mass (10^{24} kg)	1,898.6	568.46	86.832	102.43
Volume (10^{10} km^3)	143,128	82,713	6,833	6,254
Equatorial radius (km)	71,492	60,268	25,559	24,764
Ellipticity (oblateness)	0.06487	0.09796	0.02293	0.01708
Mean density (kg/m^3)	1,326	687	1,270	1,638
Surface gravity (m/s^2)	23.12	8.96	8.69	11.00
Surface temperature (Celsius)	–140	–160	–180	–200
Satellites[a]	63	60	27	13
Mean distance from Sun millions of km (miles)	779 (483)	1,434 (891)	2,872 (1,785)	4,495 (2,793)
Rotational period (hrs)[b]	9.9250	10.656	–17.24	16.11
Orbital period	11.86 yrs	29.66 yrs	84.01 yrs	164.79 yrs

Notes:
a. Numbers are for *known* satellites as of the year 2009.
b. Retrograde rotational periods are preceded by a minus sign.
Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.

oblateness of about 0.1, which means that its equatorial diameter is about 10 percent bigger than its polar diameter, and thus it appears noticeably flattened at the poles. Saturn takes 10 hours and 13 minutes to make one complete rotation; Jupiter spins even faster, taking only 9 hours and 55 minutes to complete a rotation. Uranus takes 17 hours for one complete rotation, while Neptune takes about 16 hours (a figure that remains subject to conjecture in the aftermath of Voyager 2 studies in 1989).

These rapid rotations are surprising for such gaseous and liquid planets, because an angular momentum principle of elementary physics would have bigger bodies rotate more slowly than smaller ones. Such rapid rotation rates, then, are a mystery of the first magnitude in solar physics and geophysics. The correlation between magnetic fields and rotation rates is not very strong. Why some planets have powerful magnetic fields and others negligible ones is unknown. In general, however, if magnetic fields result from dynamo currents deep within planets, then rapid rotators should have strong magnetic fields. This is largely true in the case of both Jupiter and Saturn. Uranus has a magnetic field weaker than that of Saturn. Neptune has a magnetic field roughly comparable to that of Uranus.

Neptune and Uranus seem to be smaller and colder versions of Saturn and Jupiter. Uranus has a pale blue, almost greenish-blue appearance, undoubtedly because of the presence of methane. Imaging from Voyager photographs has shown that Uranus has belts and zones, although they are not as spectacular as those of Jupiter and Saturn. Ammonia within the Uranian atmosphere and that of Neptune is thought to have precipitated to the surface. In August, 1989, Voyager 2 flew by Neptune and returned images of that planet and its atmosphere. No good photographs of Neptunian surface features can be made from Earth's surface because of its distance (30 AU from the Sun). However, the Hubble Space Telescope has been used for intermittent studies of both Uranus and Neptune, some of the most productive research being performed by astronomer Heidi Hammel. Voyager 2 showed that Neptune is also a pale blue planet with irregular marked bands in its atmosphere. It has a gigantic Dark Spot, somewhat analogous to Jupiter's Great Red Spot, which seems to cause a tremendous sinking and upswelling of its atmospheric winds. This spot has a diameter about the same as the diameter of Earth. In addition, extremely high clouds, about 50 kilometers above the (normal) Neptunian atmosphere, make its atmosphere different from those of the other Jovian planets. Winds on the order of 200 meters per second have been measured in the Neptunian atmosphere, with unique streams and band systems.

All Jovian planets possess numerous moons, or satellites. By 2008, Neptune was discovered to have at least thirteen satellites. Uranus has at least twenty-seven detected moons, while Saturn has sixty and Jupiter has at least sixty-three. It is a bit surprising that these planets rotate so rapidly while at the same time each has such a large family of satellites. The presence of such satellites should have slowed the rotation rates, if indeed the satellites and the primary planets had common origins. Many of the satellites present the same face toward their primary planet, and thus are tidally locked.

All Jovian moons have rings as well, although they vary considerably in texture and content. The Voyager 1 spacecraft discovered a very thin ring around Jupiter in 1979. Since that time research has revealed that there is more structure to Jupiter's ring system. Presently it is referred to as the Halo Ring, the Main Ring, and the Gossamer Ring; the Gossamer Ring itself has two parts, the inner Almathea Gossamer Ring and the outer Thebe Gossamer Ring. Jupiter's rings are dark and composed of dust, and therefore they cannot be seen from Earth. Saturn's ring system of mainly icy particles is incredibly dynamic, with a number of gaps due to gravitational resonances and shepherding by small, embedded moonlets. Uranus has at least nine complete rings, some considerably brighter than others; it must be noted that five of Uranus's rings were discovered in the late 1970's by ground-based observations, not by the Voyager 2 mission, although the Voyager probe provided the first intense study of the entire ring system. It was Hubble-based research that found four additional dark rings long after the Voyager flyby. Neptune has only partial arcs. Astronomers are still debating whether partial arcs are ring systems being formed or in late stages.

METHODS OF STUDY

Galileo's discovery in 1610 of the four large natural satellites of Jupiter (Io, Europa, Ganymede, and Callisto, known as the Medician and then later as the Galilean moons) launched the age of modern science. By the 1650's, Christiaan Huygens in Holland and other astronomers in Italy had established conclusively that Saturn had rings and at least one large satellite, Titan. The Great Red Spot on Jupiter has been observed ever since 1660, and the zones and belts on both Jupiter and Saturn had been clearly detected by enterprising visual astronomers. For three centuries, astronomers around the globe have tracked the Great Red Spot and noted changes in the belts and zones of these two gigantic planets. Uranus, too, was viewed by many from the 1700's onward, but it was not clearly designated as a planet until the late eighteenth century.

With the advent of spectroscopy in the nineteenth century, helium was discovered first on the Sun and shortly thereafter on Jupiter and Saturn. In the early twentieth century, it was learned that methane and ammonia were present in the Jovian atmospheres as well. In 1955, radio astronomers detected radio signals coming from Jupiter's magnetosphere.

It was not until the 1970's, however, that the greatest discoveries about the Jovian planets were made. Data from Voyager 1 revealed the unexpected existence of rings around Jupiter. Earth-based observations showed rings around Uranus and Neptune, and images returned from Voyager 2 in the 1980's revealed ten satellites circling Uranus and orbiting Neptune.

Radio astronomy probes the decimetric and decameter radio signals emitted from Jupiter, which signal the extent of its magnetosphere and the relationship of its halo and its volcanic innermost satellite Io, respectively. Radio astronomy conducted by the Cassini spacecraft in orbit about Saturn provided insight into the nature of the magnetic field on Enceladus. Infrared astronomy also has been very helpful in determining some of the features of the cold Jovian planets Uranus and Neptune. Many experimental techniques have been used to determine the size and extent of the ring systems surrounding these planets, and still there are many unanswered questions about these systems.

The atmospheres of Jupiter and Saturn have been probed with all sorts of sensitive spectrometers, yet experimental information is valid only for a penetration depth of a few tens of kilometers. What lies below the turbulent, fast-moving atmosphere has not been experimentally detected; all that scientists can do is rely on the best theories and modeling techniques presently available.

The Pioneer 10 and 11 probes found that Jupiter is a tremendous source of electrons and that it generates several times as much heat as it receives from the Sun. The origin of these electrons and heat is far from clear to the most discriminating theorists in both physics and geophysics. There are no comparable conditions on Earth or the nearby terrestrial planets to produce such effects. Voyager 2 passed Neptune in August, 1989, and its use as a planetary probe effectively ceased. Its next primary objective was to characterize the approach to interstellar space. The Hubble Space Telescope (HST) has produced much better images of Uranus and Neptune than had previously been available on a regular basis.

The Galileo spacecraft arrived at Jupiter in 1995. A special probe released from Galileo entered Jupiter's atmosphere on December 7, 1995, and its instruments detected a new radiation belt, fierce winds, lightning, and upper-atmosphere densities and temperatures much higher than expected. The Cassini spacecraft, launched in October, 1997, began exploring the Saturn system from an orbital vantage point beginning in 2004. The Huygens probe that Cassini carried along on its journey from Earth to Saturn was released and sent down through the atmosphere of Saturn's largest satellite, Titan, a moon with a thick atmosphere that obscures its surface. Cassini carried an imaging radar to map the satellite's surface during repeated close flybys. Huygens survived its plunge through the atmosphere and landed in a mushy, cryogenic surface. Huygens sent its data on two redundant channels, but, because of a software error, only one transmitted properly; fortunately, an alternative path recovered most of the data that otherwise could have been lost. Huygens and Cassini found evidence of complex hydrocarbons under cryogenic conditions on the surface of Titan. Cassini's primary mission was completed in 2008, and the program received a fully funded two-year extension.

CONTEXT

It was expected that early exploration of the Jovian planets and their extensive satellite systems would provide scientific clues as to how the

Jovian Planets' Atmospheres: Comparative Data

	Jupiter	*Saturn*	*Uranus*	*Neptune*
Surface pressure (bars)	>100	>100	>100	>100
Surface density (kg/m^3)	~0.16	~0.19	~0.42	~0.45
Avg. temperature (kelvins)	~129	~97	~58	~58
Scale height (km)	27	59.5	27.7	~20
Composition				
Ammonia	260 ppm	125 ppm	tr	tr
Ethane	5.8 ppm	7 ppm	—	1.5 ppm
Helium (%)	10.2	3.25	15.2	19
Hydrogen (%)	89.8	96.3	82.5	80
Hydrogen deuteride (ppm)	28	110	~148	192
Methane	3000 ppm	4500 ppm	~2.3%	1.5%
Water	—	~4 ppm	tr	tr

Notes: Composition: % = percentages; ppm = parts per million; tr = trace amounts.
Source: Data are from the National Space Science Data Center, NASA/Goddard Space Flight Center.

solar system formed. Instead, a whole series of new mysteries has appeared, spurring additional robotic exploration of the outer solar system.

For atmospheric physicists, the weather systems evident in the atmospheres of both Jupiter and Saturn have provided much material for study. The Great Red Spot and several of the lesser white spots on both Saturn and Jupiter have proved to be cyclonic or anticyclonic storms that are able to maintain themselves for decades. The Great Red Spot has been observed for at least for 350 years. Could such a massive storm system be maintained on Earth? What conditions on Jupiter contribute to the tremendous longevity of the Great Red Spot? In attempting to answer questions of this sort, scientists have modeled all manner of weather systems, which has proved to be useful in deciphering meteorological patterns on Earth. Thus Jupiter and Saturn have served as gigantic, high-pressure, turbulent laboratories for atmospheric modelers. Indeed, the greatest potential outcome of comparative planetology is a better understanding of complex geophysical and atmospheric physics processes right here on Earth. Such an understanding is fundamental to determining whether or not Earth is presently undergoing global warming of natural or human-made origin.

Even in the esoteric discipline of fluid mechanics, particularly in the study of turbulent flow, data from Jupiter and Saturn have been unexpectedly helpful. These studies are critical in air-frame design and, when coupled with modern computer modeling techniques, have proved to be very valuable in the design of supersonic air frames and high-speed hydrofoils. Neptune's Dark Spot should provide fodder for studies of both fluid-mechanics and meteorology well into the twenty-first century. Many scientists believe that solar-system locations most likely to host life or organic chemistry necessary for life are either Jupiter's ice-covered satellite Europa, or Saturn's satellites Titan and Enceladus. Some sort of life systems could be operating in either of these locations, for the energy and chemical conditions seem suitable. Should some sort of complex organic molecules or anaerobic bacteria be found on either Jupiter or Europa, the perennial mystery as to how life formed on Earth and why it exists at all could be addressed intelligently, perhaps for the first time. Saturn's satellite Enceladus has cryogenic geysers in its south polar region. Neptune's satellite Triton has also appeared to contain some sort of cryogenic geyser activity. The unexpected detection of warm liquids in the outer solar system could drive biological networks. Titan has organic materials that are believed to be indicative of the primordial Earth, although the satellite is far colder than Earth was when life developed here. Thus Titan might be a frozen example of what the early Earth might have been when life first arose.

In 1955 radio astronomers Bernard Burke and Kenneth Franklin, while studying the Crab nebula, inadvertently discovered radio emissions coming from Jupiter. Decades later, Voyager 2 recorded the largest electrical current ever measured as it passed near Jupiter. In the first half of the twentieth century, most scientists did not realize that Earth, with its reasonably strong magnetic field, produced a magnetosphere just as Jupiter did. It was not until early American spacecraft discovered the Van Allen radiation belts that radio engineers, astronomers, and plasma physicists realized that Earth's magnetosphere was a smaller version of Jupiter's. The magnetospheres of Jupiter, Saturn, and even Earth are still not completely understood; what influence they might have had on planetary origins and developments is unknown. Eventually, studies of Jupiter and Saturn might provide clues regarding the forces and mechanisms behind electrical storms, violent atmospheric electricity, and radio blackouts that can have pronounced effects on life on Earth.

John P. Kenny

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. Cambridge, England: Cambridge University Press, 2007. A comprehensive work about the biggest planet in the solar system, including a series of articles provided by recognized experts in their fields of study. Excellent repository of photography, diagrams, and figures about the Jupiter system and the various spacecraft missions that unveiled its secrets.

Bortolotti, Dan. *Exploring Saturn*. New York: Firefly Books, 2003. A look at the Cassini-Huygens mission for a younger audience. Full of charts, photographs, a section on observing Saturn, and an overview of the history of our understanding of the Saturnian system, from antiquity to the launch of Cassini.

Greenberg, Richard. *Europa the Ocean Moon: Search for an Alien Biosphere*. New York: Springer, 2005. A complete description of Europa through the post-Galileo spacecraft era. Discusses the astrobiological implications of an ocean underneath Europa's icy crust. Well-illustrated and readable by both astronomy enthusiasts and college students.

Harland, David H. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. New York: Springer Praxis, 2000. Collects virtually all of NASA's press releases and science updates during the first five years of the Galileo mission, along with a preview of the Cassini mission. Includes an enormous number of diagrams, tables, lists, and photographs.

Harland, David M. *Cassini at Saturn: Huygens Results*. New York: Springer, 2007. Provides a thorough explanation of the entire Cassini program, including the Huygens landing on Saturn's largest satellite. Essentially a complete collection of NASA releases from the start of Cassini flight operations through the majority of Cassini's seventy orbits during its primary mission (Cassini's primary mission concluded a year after this book was published). Technical writing style but accessible to a wide audience.

_______. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002. A volume in Springer's Space Exploration series, this is a technical description of the Cassini program, its science goals, and the instruments used to accomplish those goals. Written before Cassini arrived at Saturn. Provides a historical review of pre-Cassini knowledge of the Saturn system.

Hartmann, William K., ed. *Astronomy*. 5th ed. Belmont, Calif.: Wadsworth, 2004. Hartmann's section of this astronomy textbook, which should be accessible to high school and college students, examines the Jovian planets and other parts of the solar system. Besides discussing many late twentieth century findings, Hartmann lists various theories of planetary origins and natures, and he examines the strengths and weaknesses of each. One chapter focuses on Jupiter, and another compares Jupiter to the other Jovian planets. Well illustrated.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find the standard texts for introductory astronomy too low-level. Covers all topics from solar-system objects to cosmology.

Lovett, Laura, Joan Harvath, and Jeff Cuzzi. *Saturn: A New View*. New York: Harry N. Abrams, 2006. A coffee-table book with about 150 of the best images returned by the Cassini mission to Saturn. Covers the planet, its many satellites, and the complex ring system.

Russell, Christopher T. *The Cassini-Huygens Mission: Orbiter Remote Sensing Investigations*. New York: Springer, 2006. Provides a thorough explanation of the remote-sensing investigations of both the Cassini orbiter and the Huygens lander. Outlines the scientific objectives of all instruments on the spacecraft and describes the planned forty-four encounters with Titan. Given the publication date, only early science returns are discussed.

See also: Jupiter's Atmosphere; Jupiter's Great Red Spot; Jupiter's Interior; Jupiter's Magnetic Field and Radiation Belts; Jupiter's Ring System; Jupiter's Satellites; Neptune's Interior; Planetary Atmospheres; Planetary Formation; Solar System: Element Distribution.

Jupiter's Atmosphere

Categories: Planets and Planetology; The Jovian System

Jupiter's atmosphere differs greatly from that on Earth. It is composed mainly of hydrogen and helium and is far enough from the Sun that the temperature of the visible cloud deck is only 153 kelvins. Voyager spacecraft data revealed details concerning chemical composition, heat transport, and wind patterns within the atmosphere. The Galileo spacecraft's atmospheric entry probe sampled that atmosphere directly and forced a rethinking of the physical model of Jupiter.

OVERVIEW

Observed through an Earth-based telescope, Jupiter's most striking aspects are a pearly glow reflected from the planet; a series of east-west bands of pastel yellows, whites, browns, and blues; and the oblate aspect of its disk. The equatorial diameter is 6 percent larger than the polar diameter, a fact that is readily apparent to the observer. Closer inspection will reveal that distinctly bright and darker individual cloud features are visible within the banded structure of the atmosphere. Throughout an evening, an observer will notice that cloud features are rotating from west to east at a rate of about 36° per hour, indicating that Jupiter rotates on its axis in slightly less than ten hours. A careful student of Jupiter will discover that the planet's visible cloud deck does not rotate as a solid body; instead, eastward winds at the equator sweep the clouds past those at midlatitudes at a rate that displaces them 7° eastward each day. This type of motion indicates that the cloud deck is opaque. In order to understand the atmospheric winds, one must know how fast the core of the planet is rotating.

Radio astronomers collected data from Jupiter and realized that variations in the radio signals from Jupiter could be attributed to the interaction of the charged particles ejected from the Sun with the magnetic field of the planet. Although the signal that the radio astronomers were measuring was generated above the atmosphere of the planet, the signal varied as Jupiter's magnetic field rotated. Astronomers' understanding of magnetic fields led them to believe that the radio astronomers were measuring the rotation of Jupiter's core. They determined that one rotation took 9 hours, 55 minutes, and 29.771 seconds.

A ground-based observer equipped with an eyepiece outfitted with a crosshair can center the planet and record the time that a selected cloud feature takes to rotate past the crosshair. Because the planet rotates about its axis in approximately ten hours, while the observer is constrained to observe within a twenty-four-hour time frame, the feature will be visible on alternate nights. The rate of rotation of the feature across the visible disk of the planet, however, is such that the observer will find that the cloud rotates five times in slightly more than two days.

Accurate periods of rotation were determined by the British Astronomical Association and the American Lunar and Planetary Institute during the first half of the twentieth century. Careful measurements of photographic data by Elmer Reese from 1960 to 1974 refined these data. When the data were related to the planetary core using the radio period of rotation, an alternating pattern emerged. Strong eastward winds near the equator, as swift as 150 meters per second (more than 450 kilometers per hour), decreased poleward. Near 15° latitude, the prevailing wind in both hemispheres was westward. Between 10° and 35° latitude, the displacement of clouds revealed two westward and two eastward peaks in the horizontal wind speeds. Highly reflective regions called zones were bracketed on the equatorward side by westward winds and on the poleward side by eastward winds. The less reflective, browner regions, or belts, were nested between the zones. Horizontal wind flow of this type in Earth's atmosphere would generate conditions that would cause rising air in the zones, leading to the formation of ice clouds at high altitudes. Air would descend in the belts and cause ices to melt, allowing a longer line-of-sight through the atmosphere and more absorption of light—hence, less reflection. Recognition of this general circulation pattern in the 1960's led to questions concerning the nature of Jupiter's ices.

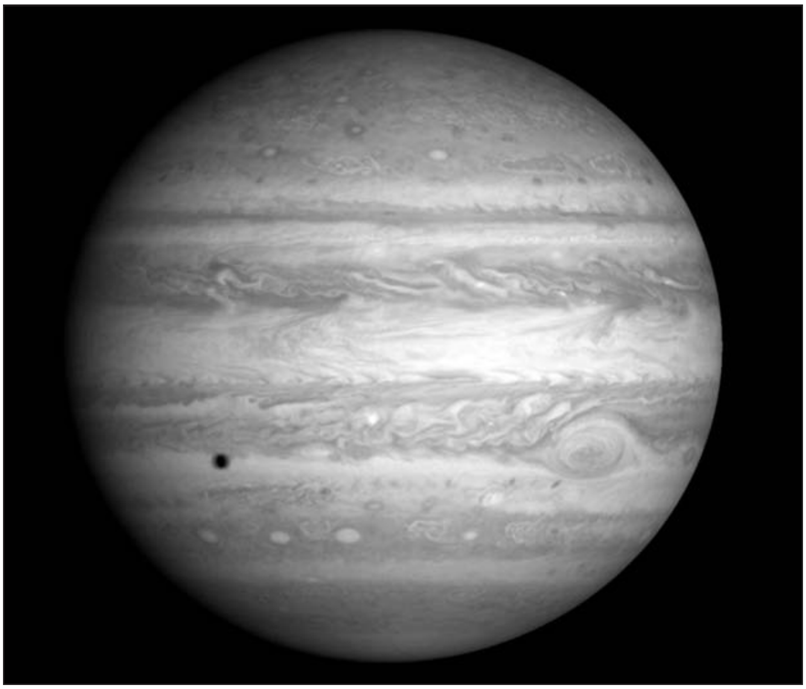
As light travels outward from the Sun, it spreads out equally; thus, its ability to heat a surface decreases rapidly via an inverse square relationship in all radial directions. By the time sunlight reaches Jupiter, at a distance five times greater than Earth's distance from the Sun, the intensity is diluted by a factor of twenty-five. This dilution leads to temperatures too low to allow melting of ice formed from water; therefore, the visible cloud deck must contain another kind of ice.

In an effort to refine their understanding of the temperature regime, astronomers began to make infrared measurements of Jupiter's atmosphere. They determined that the planet radiates one and a half times more heat than it absorbs from the Sun. These results imply that the interior of Jupiter is hotter than the cloud deck and that convection from the interior transports heat outward. The picture of a deep atmosphere dominated by east-west winds emerged. If Jupiter were composed of the same chemical mixture as the Sun, its atmospheric gas would be so strongly compressed that deep below the visible cloud deck it would form a sea of liquid hydrogen and helium. There would be an indistinct change between the surface of the cryogenic fluids and the atmosphere. It also became apparent that it was essential to understand the chemistry of the atmosphere.

Calculations carried out by John Lewis and Ronald Prinn led to a model of the atmosphere that posited the existence of an upper cloud layer of ammonia ice, underlain by an ammonium hydrosulfide layer and a cloud layer composed of water ice (where the pressure is greater than ten times Earth's atmospheric pressure at sea level). All these ices are white; therefore the calculations yielded no information about coloring agents in the Jovian atmosphere. Above the topmost cloud deck, the atmosphere is mainly hydrogen and helium gas, with traces of ammonia and methane.

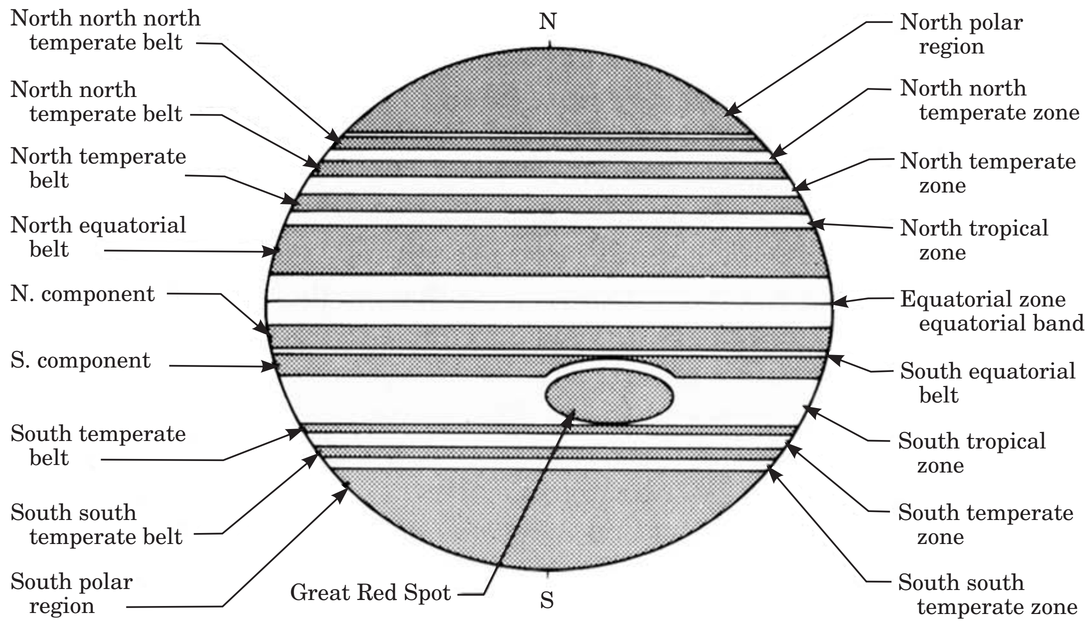
Molecules of methane and ammonia can absorb enough energy from incident ultraviolet light to break bonds, which allows the hydrogen atoms to escape. Darrel Strobel performed calculations indicating that ionized molecules combine to form more complex molecules and possibly aerosols or hazes. Laboratory work by other investigators has shown that many of the compounds that are formed are yellow and brown. Much of the color variation in the Jovian atmosphere may arise from variations in the heights of the underlying clouds.

Removal of the smog occurs in two ways. Either the particles grow large and fall to lower levels, or convective clouds of ammonia are carried upward like thunderheads, and the ammonia ice encases the smog particles and again causes them to fall to lower levels in the atmosphere. Computer modeling of scattering and transmission of light through the layers of hazes, gas, and clouds, carried out by Martin



The Cassini spacecraft delivered four images in December, 2000, that were compiled to produce this picture of Jupiter with a resolution of 144 kilometers. (NASA/JPL/University of Arizona)

Major Features of Jupiter



Source: Morrison, David, and Jane Samz. *Voyage to Jupiter*. NASA SP-439. Washington, D.C.: National Aeronautics and Space Administration, 1980, p. 4.

Tomasko, Robert West, and others, supports the theory that colorization is dependent on the varying heights of underlying clouds.

These calculations cannot explain Jupiter's Great Red Spot—neither its colors nor its long life. This feature is the largest Jovian cloud system. North to south, it spans a distance slightly larger than the diameter of Earth, and it extends farther than two Earth diameters in the east-west direction. This large, unique cloud system is trapped between a westward wind on the equatorward side and an eastward wind on the poleward side. The winds are diverted around its perimeter, rotating it in a counterclockwise direction. In Earth's atmosphere, a weather system with similar motion would rise in the center, spiral outward at the top of the cloud deck, and descend around its perimeter. The degree of redness of the Spot varies with time. A unique property of the Red Spot is its ability to absorb ultraviolet, violet, and blue light. Apparently some constituent that has been carried up from lower, warmer depths absorbs the blue light, causing the Spot to appear redder than the surrounding clouds.

Observations of Jupiter in the infrared indicate that brown and blue-gray regions are warmer than white areas. Thus, the white zones are colder than the brown belts. The clouds above the Red Spot are cold, but there is a warmer region around its perimeter. The general heat loss from the planet indicates that the interior is warmer than the upper cloud deck that radiates to space; hence, infrared maps allow astronomers to determine relative heights of clouds. A desire to obtain high-resolution maps of the infrared data and the structure of the cloud deck led to the development of the instruments that were on board the Pioneer and Voyager missions. Perhaps the best way to determine a planet's atmospheric character is to directly encounter it. That was the reasoning behind having the Galileo spacecraft carry an atmospheric probe that then was dispatched to hit the upper atmosphere and relay data until tremendous pressures destroyed it.

KNOWLEDGE GAINED

The Pioneer 10 and 11 spacecraft, which arrived at Jupiter in November, 1973, and November, 1974, respectively, each carried three instruments that sampled the Jovian atmosphere. The fact that the Pioneer spacecraft were spin-stabilized limited the types of instruments that could be placed on board. Voyagers 1 and 2 passed through the Jupiter system four months apart in 1979. Ultraviolet and infrared instruments and a spin-scan camera were trained on Jupiter's atmosphere. The Galileo probe also sampled the atmosphere beginning in 1995. The Cassini spacecraft in December, 2000, and the New Horizons spacecraft in February, 2007, conducted in-depth studies of Jupiter as they flew past on their way to Saturn and Pluto, respectively.

Infrared data revealed that there was little temperature variation between the equator and the pole at the cloud-top level. Data indicated there are limits to the role that equatorial solar heating plays in driving the zonal winds. Andrew Ingersoll proposed that solar heating at the equator could bring about a cloud structure that would act as an insulating blanket, causing the heat from the interior to emerge near the poles and resulting in little temperature variation at the level of the ammonia cloud deck. This hypothesis implies that the outward transport of the heat from the interior may dominate the atmospheric wind patterns.

Pioneer spin-scan cameras were equipped with blue and red filters and polarizers. The nature of the camera did not allow a large number of images to be obtained. Nevertheless, data provided valuable material for study of the scattering properties of the atmospheric smog and haze layers. Although a series of images with sufficiently high resolution to map cloud motions could not be obtained, images of the Red Spot and north polar regions confirmed information previously gained from ground-based observation. They also provided data on the scale of the cloud structures.

Voyagers 1 and 2 carried five instruments that were used to observe Jupiter's atmosphere: two television cameras (one with a wide-angle view and the other with higher resolution and a narrow field of view), infrared and ultraviolet spectrographs, and a photopolarimeter. Multicolor high-resolution mapping of the visible cloud deck could be obtained at three-month intervals with each spacecraft. Near-encounter infrared measurements resolved temperature variations as a function of latitude and longitude on the planet. The ultraviolet spectrometer obtained data concerning high altitudes in the Jovian atmosphere. This extensive data set has been combined with the Pioneer and historical ground-based data sets in an effort to shed light on both short-term and long-term atmospheric variations. Cloud displacements were measured by Reta Beebe, Ingersoll, and others. Eastward winds near the equator were as powerful as 160 meters per second. Westward wind speeds at 15° north and 17.5° south latitude were both retrograde at 40 and 70 meters per second, respectively. Eastward wind maxima at 20° north and 24.5° south latitude were 170 and 60 meters per second, respectively. Voyager scientists found considerable differences between the magnitudes of wind jets in the northern hemisphere and those in the southern hemisphere; no change in the average zonal wind was detected at any latitude, however, during the five-month interval between the two encounters.

Infrared measurements indicated that the winds decrease with height above the deck, and that temperatures and abundance of ammonia above the cloud deck are consistent with an atmosphere that is driven by cloud motions at the level of the visible cloud deck.

Galileo's entry probe hit Jupiter's upper atmosphere on December 7, 1995, at a speed of as much as 170,000 kilometers per hour. The atmosphere decelerated the probe at an increased g-load of approximately 230 (that means the force was 230 times that of normal Earth gravity at sea level). During the probe's 57-minute-long plunge, it successfully relayed its findings to the Galileo orbiter for storage and eventual playback to Earth. Galileo was a little over 200,000 kilometers above the probe at the time.

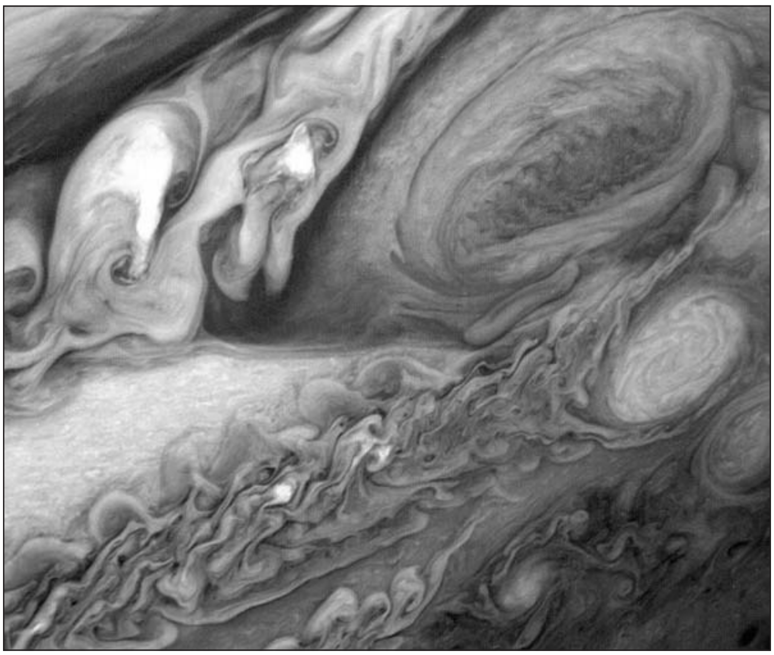
The probe provided some surprising data. It had been hoped that the probe would find considerable amounts of water vapor in the atmosphere and detect extensive amounts of electrical activity, or lighting. In reality it found very little of either. It did find a new radiation belt just 50,000 kilometers above the cloud tops. As it descended through Jupiter's atmosphere, the probe registered very strong winds and experienced significant turbulence. Spectrometers found lower abundances of helium, neon, carbon, oxygen, and sulfur than had been expected. Helium was nearly half as abundant as expected in contemporary atmospheric models for Jupiter; Galileo researchers were expecting the probe to fall through a three-layered cloud structure. The probe did not experience anything like what had been predicted. The net flux radiometer on the probe did find some high-level ammonia ice clouds, and the nephelometer instrument provided some evidence of ammonium hydrosulfide clouds. Water ice was absent, suggesting the probe had entered one of the driest spots on Jupiter.

Wind strengths and atmospheric temperatures varied during the probe's descent. Winds reached 350 kilometers per hour with gusts up to 525 kilometers per hour. After the probe had plunged 156 kilometers through the atmosphere under its main parachute, the temperature and pressure environment destroyed it; most likely it was crushed, vaporized, or both nearly simultaneously. Essentially the probe's encounter forced planetary scientists to rethink the current model of Jupiter's atmosphere.

CONTEXT

By the mid-nineteenth century, astronomers had become aware that Jupiter was unlike Earth. Using the basic laws of motion, the apparent size of Jupiter, and the known distances within the solar system, they determined that even though the volume of Jupiter was more than eleven hundred times larger than Earth's volume, its mass was only 318 times larger than that of Earth. Thus, although this planet is much more massive than Earth, gravity has not compressed Jupiter's interior to the high densities that are present in the interior of Earth. Nineteenth century astronomers concluded that Jupiter could not have the same chemical composition as Earth.

By 1960, spectroscopists had determined that the atmosphere of Jupiter was cold and that temperatures at the level of the visible clouds were near 153 kelvins. Spectra revealed absorption by molecules of methane and ammonia. These observations are consistent with an atmosphere that is composed mainly of hydrogen and helium with small amounts of carbon and nitrogen. At the observed temperatures, oxygen would combine with hydrogen to form water, which would be trapped below the visible cloud deck. It became apparent that Jupiter was composed of a chemical mixture similar to that of the Sun, and that the small silicon- and iron-rich planets of the inner solar system were very different from the outer gas-rich planets.



Voyager 1 captured this image of Jupiter's roiling atmosphere in 1979. (NASA/JPL)

Current models concerning the formation of a solar system propose that planets the size of Jupiter form first at distances far enough from the parent star that hydrogen and helium have not been expelled by the radiation from the star. The turbulence that this generates in the preplanetary gas and dust cloud leads to the formation of the other planets, with the inner ones forming from hydrogen-poor material. The importance of Jupiter-sized bodies in the formation of other planets that could possibly support other lifeforms has stimulated interest in learning more about the nature of this gas giant. However, it must be noted that extrasolar planets have been found to have masses in excess of Jupiter and to be located extremely close to their stars.

Jupiter's atmosphere is chemically unlike that of Earth. The planet's huge depth and lack of irregular landmasses at its lower boundary contrast with conditions in Earth's atmosphere. There are, however, some similarities: The main constituents of Jupiter's atmosphere are hydrogen and helium. Like the nitrogen and oxygen molecules of Earth's atmosphere, these particles do not absorb sunlight readily. A large portion of solar energy passes through the upper atmospheres of these planets and, in the case of Earth, the surface absorbs the energy and is warmed. The atmosphere is then heated from the bottom, with trace constituents, carbon dioxide, and water absorbing and reradiating the energy. This leads to decreasing temperatures at increasing altitudes in the lower atmosphere.

Voyager infrared data confirmed that Jupiter has an internal heat source and that it emits 1.67 times more energy than it absorbs from the Sun. Infrared data indicate that the winds of Jupiter are driven, like those on Earth, by energy input in the lower atmosphere. Knowledge of planetary atmospheres has become more general and efforts to define the factors that lead to climate variations can be applied to more than one planet. Jupiter's dissimilarity to Earth provides checks and challenges in the search to understand Earth and the solar system.

Reta Beebe

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. Cambridge, England: Cambridge University Press, 2007. A comprehensive work about the biggest planet in the solar system, covered in a series of articles provided by recognized experts in their fields of study. Excellent repository of photography, diagrams, and figures about the Jupiter system and the various spacecraft missions that unveiled its secrets.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. Filled with color diagrams and photographs, this popular work covers solar-system astronomy and planetary exploration through the Galileo missions. Accessible to the astronomy enthusiast.

Cole, Michael D. *Galileo Spacecraft: Mission to Jupiter*. New York: Enslow, 1999. Provides a full description of the Galileo spacecraft, its mission objectives, and science returns through the primary mission. Particularly good at describing mission objectives and goals. Suitable for a younger audience.

Fimmel, Richard O., James Van Allen, and Eric Burgess. *Pioneer: First to Jupiter, Saturn, and Beyond*. NASA SP-446. Washington, D.C.: Government Printing Office, 1980. A detailed review of the original Pioneer mission. Reproduces most of the images obtained by the spin-scan camera. Suitable for the general reader.

Gehrels, Tom, ed. *Jupiter*. Tucson: University of Arizona Press, 1976. A historic collection of scientific essays covering all aspects of Jupiter. The volume reflects the state of knowledge of the planet before the Voyager mission; a subsequent survey has not been published. Its ample documentation will, however, direct the serious student to journals and other sources that update the information available here. For the advanced reader.

Harland, David H. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. New York: Springer Praxis, 2000. This book provides virtually all of NASA's press releases and science updates during the first five years of the Galileo mission. Provides a preview of the Cassini mission. Includes an enormous number of diagrams, tables, lists, and photographs. The book's description ends before completion of the Galileo mission unfortunately, but what is missing can easily be found on numerous NASA Web sites.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. The chapter on Jupiter thoroughly addresses the Jovian system and spacecraft exploration of it.

Hunt, Garry E., and Patrick Moore. *Jupiter*. New York: Rand McNally, 1981. Reviews the original Voyager mission and describes the Jovian system. Photographs and illustrations are plentiful. Requires some background knowledge.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Suitable as a textbook for upper-level college courses in planetary science.

McAnally, John W. *Jupiter, and How to Observe It*. New York: Springer, 2008. An observing guide for the amateur astronomer that also provides detailed descriptions of the Jovian system. Discusses observational techniques, including a wide range of popular telescopes and ancillary equipment.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar-system astronomy suitable for an introductory college course. Accessible to nonspecialists as well. Filled with supplemental learning aids and solved student exercises. A companion Web site is available for educator support.

Morrison, David, and Jane Samz. *Voyager to Jupiter*. NASA SP-439. Washington, D.C.: Government Printing Office, 1980. A description of the events surrounding the Voyager missions to Jupiter. Illustrated with many color photographs. Accessible to the general reader.

Peek, Bertrand M. *The Planet Jupiter*. London: Macmillan, 1958. A detailed summary of ground-based observations recorded by the British Astronomical Association. This classic book provides an overview of the time-dependent aspects of the Jovian cloud deck and the history of the Red Spot. Although dated, this accessible presentation of the basics can be compared with more contemporary understandings.

See also: Auroras; Brown Dwarfs; Comet Shoemaker-Levy 9; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Eclipses; Europa; Extrasolar Planets; Io; Jovian Planets; Jupiter's Great Red Spot; Jupiter's Interior; Jupiter's Magnetic Field and Radiation Belts; Jupiter's Ring System; Jupiter's Satellites; Planetary Atmospheres; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Saturn's Atmosphere.

Jupiter's Great Red Spot

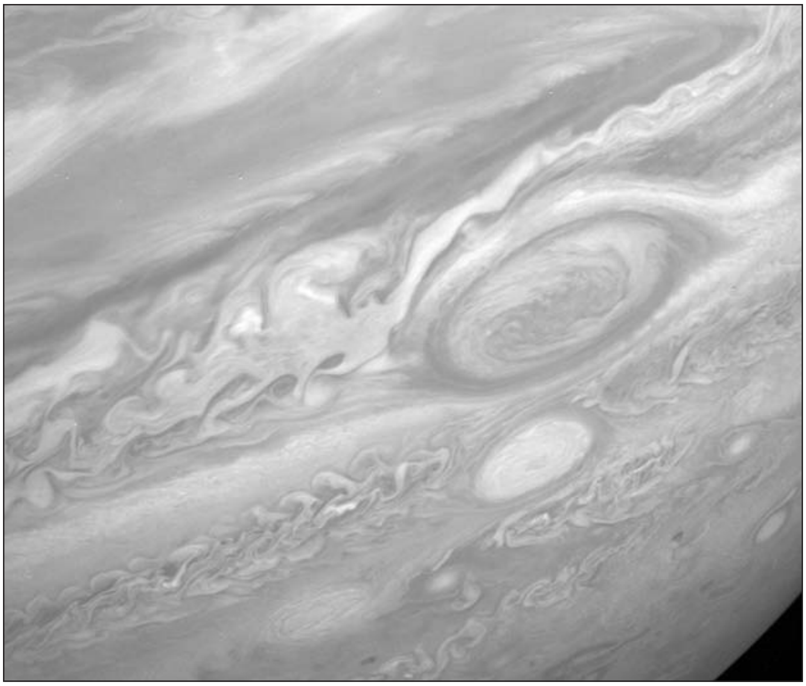
Categories: Planets and Planetology; The Jovian System

High-resolution data have been obtained concerning the nature of Jupiter's Red Spot, a weather system with horizontal dimensions comparable to the diameter of Earth and monitored behavior spanning centuries.

OVERVIEW

Excluding Jupiter's general east-west belt and zone pattern, the Great Red Spot is the most obvious, persistent, and continuously observed feature of Jupiter's visible cloud deck. Centered at about 20° south latitude, it spans about 14,000 kilometers in the north-south direction and about 26,000 kilometers in the east-west direction. Compared with the diameter of Earth, the Red Spot is a huge cloud structure large enough to span two Earths.

This well-defined oval feature has raised considerable curiosity ever since its discovery, which many credit to original observations made by Giovanni Cassini or Robert Hooke in the seventeenth century. Coordinated reports by the British Astronomical Society and original drawings maintained in the Royal Astronomical Society library collections in London definitively established that the Red Spot has been present since at least 1830. Earlier scattered reports of pink spots in the atmosphere of Jupiter extend back several centuries to the era of earliest efforts to improve the resolution of simple telescopes.



The Great Red Spot is seen in this Voyager 2 image, appearing as the large, eye-shaped oval to the right of center. (NASA/JPL)

Realizing that they were seeing an opaque cloud deck and that the surface of Jupiter was never visible, some observers suggested that the Red Spot was caused by interference between an elevated surface feature and the prevailing winds. In the early 1960's Raymond Hide proposed a model for the driving mechanism. If this model had withstood scrutiny, it would have permitted scientists to calculate the planet's surface rotation rate and to interpret other cloud motions by extrapolating from this rate. Careful examination of measured periods of rotation indicated, however, that Hide's model was inconsistent with available data. The Red Spot circles Jupiter at an almost unvarying speed for a period of twenty to fifty years. At the end of each period, it suffers an acceleration or deceleration which occurs over a period of weeks. After this adjustment period, the Red Spot continues to circle at its new speed. This behavior indicates that it cannot be the result of an upwardly propagating disturbance above an elevated region on a hidden surface. Measured positions indicate that there is no rate at which the interior of the planet could rotate that would not force the Red Spot to drift freely either east or west within the atmosphere, yet historical observations indicate that the spot is trapped in the prevailing east-west wind pattern and is not free to move north or south like weather systems in Earth's midlatitudinal regions.

Along the southern edge of Jupiter's equatorial zone, winds blow eastward at 150 meters per second. West-to-east zonal winds decrease poleward, until at 17.5° south latitude they are moving westward at 70 meters per second. From 17.5° to 24.5° latitude, winds increase eastward to a maximum of nearly 60 meters per second. From 24.5° to 50°, winds alternate eastward and westward. This alternating east-west wind pattern, with four cycles between the equator and 50° south latitude, generates significant latitudinal wind shear. If local heating occurs below the cloud deck, causing the atmosphere to rise and clouds to form, maintaining a long-lived cloud system in the presence of strong horizontal shear would require the cloud to rotate about its center. If the cloud rotates in the same sense as the local horizontal shear, it can deflect the prevailing winds about its perimeter. The Red Spot displays this behavior. Not only is it trapped between westward winds at 17.5° and eastward winds at 24.5°, it also deflects westward wind flow around its equatorward perimeter, creating a large indentation, or hollow, in the poleward side of the dark adjacent belt. Other, smaller oval cloud systems are associated with the more poleward wind-shear regions. Three white oval cloud systems, noted in 1938, are located near 29° south latitude. The east-west dimension of each of these systems is about 12,000 kilometers. A series of smaller ovals circle the planet near 37° south latitude. Morphologically, the Red Spot is not unique. However, it is the largest example of a type of cloud system common to the southern hemisphere of the planet.

The Red Spot is notable not only in size but also in coloration. Jupiter's other oval clouds are white, indicating that their cloud decks are composed of highly reflective ammonia ices. When visible red and infrared reflection from the Red Spot is analyzed, data indicate ammonia ice is present there as well. The Red Spot has additional trace constituents in its cloud deck that are strong absorbers of ultraviolet, violet, and blue wavelengths. This gives the Red Spot its unique color. Small, short-lived ovals that form at similar latitudes in the northern hemisphere also absorb ultraviolet and blue light. This suggests that these absorbers are carried upward from below. Also the rate of vertical motion or the depth to which the convective motion reaches permits transport not present at the top of the cloud deck in storms located at more poleward latitudes.

A trip to a mountaintop on Earth's surface makes it clear that lower elevations of Earth's atmosphere are compressed. Jupiter's atmosphere must behave similarly. In order for the Red Spot to behave as an isolated system, its vertical dimension must be small in relation to its horizontal extent. Comparisons of the Red Spot with a hurricane are inappropriate. The Red Spot is a giant rotating cloud system, trapped in the prevailing winds. Reflectivity and degree of redness vary with time; still, deflection of the westward jet around the equatorward side of the Red Spot is always visible.

In 1878 the Red Spot suffered a deceleration. The surrounding cloud deck became highly reflective and white; however, the Red Spot remained dark and red. This sharp contrast made many casual observers aware of the phenomenon. In 1901, a disturbance occurred in the South Tropical Zone, the white band south of the Red Spot. This event appeared to be a major weather disturbance that moved eastward and caught up with and then passed the Red Spot, thereby accelerating the Spot. This continued until 1938. Then the belt just south of the South Tropical Zone underwent a major disruption resulting in greatly increased reflectivity of the belt and the formation of three white ovals. In the early 1930's, the Red Spot drifted at a rate similar to that seen prior to 1878. After formation of these ovals, the Red Spot decelerated to its slowest drift rate ever observed. Since 1962, the Red Spot has been drifting at a rate similar to that of the 1878-1901 period.

This constitutes evidence that the Red Spot interacts with its surroundings and that variations in local temperature, pressure, and wind patterns occur. Even so, the entire range of variation in average Red Spot motion, with the average velocity derived from the annual longitudinal displacement of the Spot relative to the rotation rate of radio noise, lies between –4.4 and –0.6 meters per second. Although this variation is small when compared to daily wind speeds at midlatitudes on Earth, an annual increase of 2 meters per second in wind speed results in an eastward displacement of about 63,000 kilometers, or about two and a half times the Spot's length.

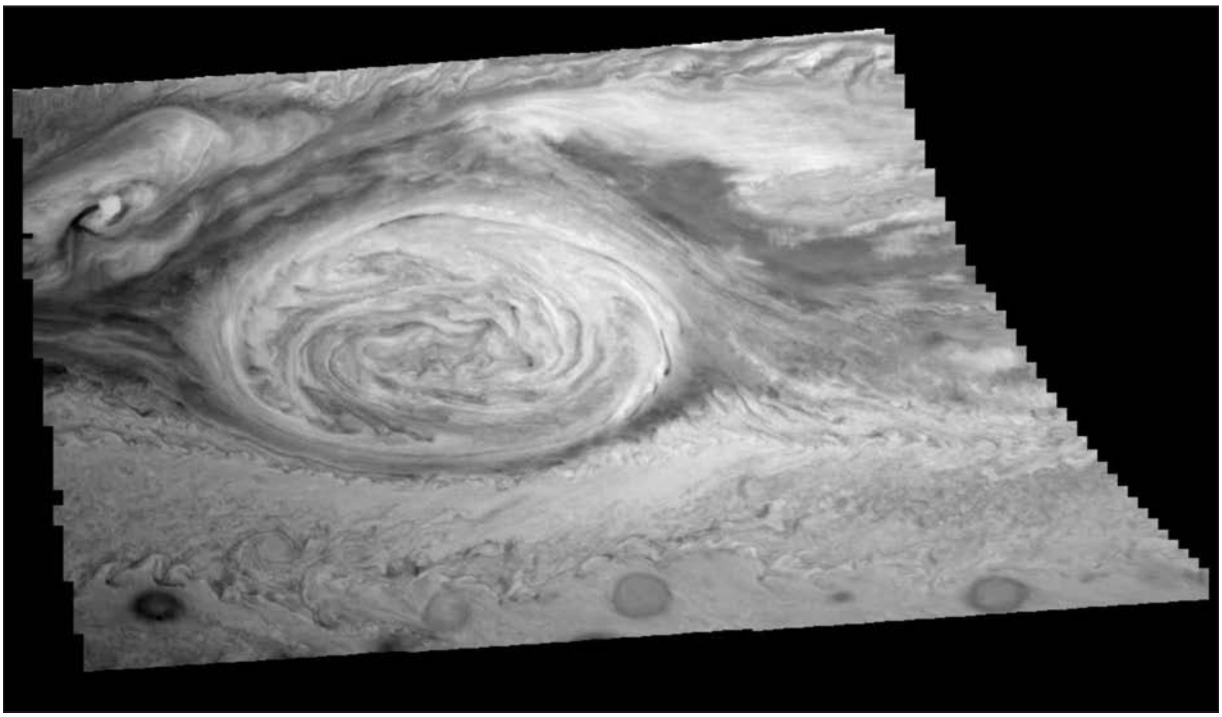
That the Spot's recovery from a given acceleration or deceleration takes years is expected. A body's heat loss rate depends heavily on the relative temperatures of the body and its surroundings. The Jovian cloud deck temperature is approximately 153 kelvins. Thus, the rate of heat loss to black sky is relatively slow. It is logical that once an excess amount of heat has been inserted into the atmosphere, it will be several years before the atmosphere returns to its previous state. One basic question that atmospheric scientists wanted to answer concerned the nature of acceleration mechanisms. Ground-based observations indicated that these events occurred over short time intervals.

During the 1960's and early 1970's, Elmer Reese made many detailed measurements using photographs of Jupiter. One result of this work was a measurement of Red Spot rotation.

The Red Spot rotates counterclockwise, completing one rotation every twelve days. A feature in Earth's atmosphere with this behavior would have air rising in the center, flowing outward at the cloud top, and descending around the perimeter. Measurement of divergent flow was one goal of the two Voyager spacecraft. Superimposed on the drift is an oscillatory motion of the whole feature, speeding up and slowing down so that its velocity oscillates every ninety days, causing the Spot to shift back and forth about 900 kilometers relative to its average path. This behavior apparently results from some natural period of response of the system to its surroundings.

Because the Red Spot had already been subjected to much scrutiny, planetary scientists eagerly looked forward in turn to obtaining high-resolution data from Pioneer and Voyager flybys in the 1970's and subsequently from the Galileo orbiter between December, 1995, and September, 2003. The Galileo spacecraft's science mission centered on investigations of the planet's particles and fields, various satellites, and general atmospheric dynamics; the latter included an atmospheric probe, but that heavily instrumented payload was not flown into the Red Spot. Additional high-resolution images of the Red Spot were obtained from the Earth-orbiting Hubble Space Telescope and during flybys of probes passing through the Jovian system to gain a gravity asset from Jupiter. Those flybys provided opportunities for controllers to test scientific packages on those spacecraft. The Cassini orbiter in December, 2000, on its way to the Saturn system, obtained images of the Red Spot superior to those from the Voyager spacecraft, even though its closest approach to Jupiter was considerably farther out. New Horizons in February, 2007, on its way to the Pluto-Charon system, obtained high-resolution images of Jupiter. It also verified that observations from ground-based telescopes and the Hubble Space Telescope had indeed recorded a lightening phase of the Red Spot beginning in 2006. In addition, New Horizons studied a relatively new storm farther south of the Great Red Spot, one referred to as the Little Red Spot, a much smaller version of its long-lived sibling.

The Little Red Spot started as one of three



A near-infrared mosaic of Jupiter's Great Red Spot from images taken by the Galileo spacecraft. (NASA/JPL) white storms that formed in the 1940's. Two of those merged in 1998, and the resulting white spot merged with the third in 2000. This feature then demonstrated a trend toward reddening in late 2005, and after 2006 it was referred to as the Little Red Spot. This storm continued to grow in wind speed and in reddish hue, providing a marvelous opportunity for in-depth study by the contemporary technology of the passing New Horizons spacecraft. Comparisons of the Little and Great red spots were expected to shed light on the complex atmospheric physics raging within these large-scale and unique storms. That research continues.

KNOWLEDGE GAINED

Both Pioneer spacecraft were spin-stabilized, so that the planet swept past their instruments' fields of view. This design feature placed strong constraints on the type of instrumentation that could be implemented. The imaging experiment utilized a scanning camera that sampled one point on the planet at a time. Images were constructed by scanning the planet row by row as it passed the field of view. This method limited the ultimate resolution of the data and severely curtailed the number of images that could be recorded. Nevertheless, Pioneer data were highly useful to astronomers. Tom Gehrels and his team observed Jupiter at a time when the belt adjacent to the Red Spot was highly reflective and white. The Red Spot was quite dark. Comparison with descriptions of the Red Spot in 1878 indicated that reflectivity of the Spot and its surroundings at the time of both Pioneer encounters was highly similar to its condition approximately a century earlier.

That Red Spot behavior continued until July, 1975, when a bright white cloud appeared west of the Spot that expanded rapidly and sheared out in the zonal winds. Considerable turbulence accompanied this event, and within a few weeks the Red Spot and belt had changed significantly. Material from the disturbance encountered the Red Spot from the west and formed a large white mass of clouds to the west of the Spot. Turbulent cloud masses also spilled into the westward wind jet along the south side of the belt. This material was carried around the planet and approached the Red Spot from the east side. The contrast of the Red Spot decreased as it became whiter. Historically, Red Spot lightening has been fairly common. It appears that increased turbulence and vertical mixing in the belt that lies to the equatorward side of the Red Spot carry ammonia ices into the Spot. The Spot retained this appearance from 1975 to 1987 and was observed at high resolution in this condition by the Voyager spacecraft.

Voyagers 1 and 2 arrived at Jupiter in March and July of 1979. The two spacecraft were equipped with two television cameras each; one had a wide field of view and the other focused on a smaller field with higher resolution. The two were boresighted; thus, simultaneous views allowed detailed sampling, while defining the direction that the cameras were pointing. Ultraviolet and infrared spectrometers and a photopolarimeter allowed observations as a function of wavelength. Red Spot images with higher spatial resolution than could be obtained from Earth were taken over a period of three months with each spacecraft. The Galileo probe began collecting visible light and near-infrared images in 1996. The Hubble Space Telescope, Cassini orbiter, and New Horizons spacecraft also supplied data.

Ultimately, high-resolution sequences requiring as many as twenty-seven narrow-angle camera frames to map the Red Spot were executed. Resulting data revealed details of the flow pattern around the Red Spot. Winds were deflected around the Spot; small ammonia ice clouds, however, were observed to pass around the equatorward edge and to continue around the western cusp and along the feature's southern edge. When these clouds reached the Spot's southeast corner, they moved into the Red Spot and were sheared apart to form a high-velocity collar inside the Red Spot. Jim Mitchell, Reta Beebe, Andrew Ingersoll, and others analyzed the flow within the Red Spot and the white ovals. Velocities of rotation about the Spot's center as high as 150 meters per second were measured in the outer third of the feature. In the inner half of the Spot, reflectivity was lower, and motion of the cloud deck was small and random. No outward flow from the center toward the perimeter of the Spot was detected. When infrared data from the Red Spot and one of the white ovals were compared, no difference in absorption as a function of color could be detected; thus, the infrared data offered no clues to the identity of the ultraviolet absorber. This finding was not unexpected, because it was known that the ammonia ice would tend to dominate in the infrared.

CONTEXT

High-resolution spacecraft imagery has been combined with long-term, lower-resolution ground-based photography in an effort to understand the Red Spot's nature. Apparent motion of planetary atmospheric features can be attributed to mass motion when material is physically translated in the zonal wind or to wave motion. In the case of wave motion, variations in local pressure and temperature introduced by the wave cause local condensation or evaporation. Many of the small-scale patterns that add beauty to Earth's water clouds are of the second type. Thus, in an effort to elucidate the Red Spot's nature, models that consider different types of wave structures have been constructed. Not all wave structures are a series of oscillations with equal amplitudes traveling through space. By varying modeled environmental conditions within which the wave is formed, various researchers, including Tony Maxworthy, Andrew Ingersoll, and Gareth Williams, have investigated the characteristics of waves and related them to the morphology of the Red Spot.

In order to construct a realistic model of the Red Spot, information concerning the manner with which the zonal winds change with depth as well as with latitude is necessary. Because all required parameters are not available, a series of models must be constructed. Peter Read and others have attempted to shed light on atmospheric flow around the Red Spot by constructing cylindrical tanks of rotating fluids, within which they generate closed eddies that have characteristics similar to those of the Red Spot. High-resolution spacecraft data have provided astronomers with a wealth of information. Data have stimulated computer analysis and the gathering of additional understanding through a great deal of observation and experimentation.

It is not clear why the well-formed oval clouds in Jupiter's atmosphere preferentially form in the southern hemisphere. The fact that they are very long-lived is, however, consistent with their being large, closed eddies rotating in the local wind shear. Little is known concerning the rate of vertical motion associated with these features or the depth to which they extend below the cloud deck. Interplanetary spacecraft will continue to provide high-resolution data for researchers struggling to define the nature of Jupiter's Great Red Spot. In the planning phase for perhaps the second decade of the twenty-first century, a proposed Jupiter Icy Moons Orbiter would also provide prolonged observation of Jupiter's atmosphere, including the dynamics and evolution of the Great Red Spot.

The variability of Jupiter's atmosphere, as well as its ability to sustain prolonged features, was illustrated when a number of smaller red spots broke out beginning in early 2006. The first small spot continued into 2008, when in May a third spot appeared, this one located in the southern hemisphere farther south in longitude than the Great Red Spot. Both little red spots moved in a way that led scientists to expect them to merge. However, the influence of the Great Red Spot held the potential to push both little red spots to the side. Coordinated observations of these small storms from the Hubble Space Telescope and ground-based telescopes in visible and near-infrared light suggested that these little red spots started as white storms and then assumed a reddish color. Jupiter appeared to be undergoing a global climate change in which the equator was warming and the south pole was cooling. As a result, jet streams in the southern hemisphere were destabilized such that new storms could be generated.

Near the end of May, 2008, the first of two little Red Spots in the planet's southern hemisphere began to grow both in size and in wind speed. This new rival to the Great Red Spot developed winds of 172 meters per second, very nearly the same speed as winds in the Great Red Spot. It remained to be seen if this Little Red Spot would continue to grow independently or be swallowed up by the extremely long-lived Great Red Spot. Its general motion after its increase in size and wind speed was toward the Great Red Spot.

Reta Beebe and David G. Fisher

FURTHER READING

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. The editors have assembled a collection of articles written by top researchers in the field as a survey of then-current knowledge of the solar system bodies. One chapter places the Red Spot in its atmospheric context. Includes many illustrations.

Consolmagno, Guy. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994. A text accessible to college-level science students, using low-level mathematics as well as integral calculus where required. Demonstrates how the area of planetary science progresses by questioning previous understandings in the light of new observations.

Fimmel, Richard O., James Van Allen, and Eric Burgess. *Pioneer: First to Jupiter, Saturn, and Beyond*. NASA SP-446. Washington, D.C.: Government Printing Office, 1980. This detailed overview of the Pioneer mission includes reproductions of the Red Spot images obtained by the spin-scan camera.

Fischer, Daniel. *Mission Jupiter: The Spectacular Journey of the Galileo Spacecraft*. New York: Copernicus Books, 2001. A detailed overview of Galileo's investigations of the Jupiter system. Provides numerous full-color images taken from Galileo and other spacecraft.

Gehrels, Tom, ed. *Jupiter*. Tucson: University of Arizona Press, 1976. This collection of scientific essays on all aspects of Jupiter reflects the pre-Voyager state of knowledge of the planet. Still, its copious documentation will point the interested reader to the journals and other sources that present later findings. A challenging text; requires some background in astronomy.

Harland, David M. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. New York: Springer, 2000. Provides a detailed account of all major Galileo spacecraft operations and the mission's scientific investigations of the Jupiter system. Filled with technical diagrams and images from Galileo and other spacecraft.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An excellent text for a course on planetary science, accessible to advanced high school students and undergraduates alike. Covers the entire solar system and includes much about the Red Spot. Takes a comparative planetology approach rather than including separate chapters on individual planets in the solar system.

Hunt, Garry E., and Patrick Moore. *Jupiter*. New York: Rand McNally, 1981. Reviews the Voyager mission and describes the Jovian system. Features numerous color photographs and illustrations of the Red Spot. Accessible to the general reader.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, Neptune, and their satellites, rings, and magnetic fields. Filled with figures and photographs.

Morrison, David, and Jane Samz. *Voyager to Jupiter*. NASA SP-439. Washington, D.C.: Government Printing Office, 1980. A description of the events leading up to the Voyager missions. Communicates the excitement experienced by scientists as they received close-up views of the Red Spot. Contains many color reproductions.

Peek, Bertrand M. *The Planet Jupiter*. London: Macmillan, 1958. A classic detailed summary of ground-based observations recorded by members of the British Astronomical Association. This book provides an overview of time-dependent aspects of the behavior of the Red Spot.

See also: Comet Shoemaker-Levy 9; Jovian Planets; Jupiter's Atmosphere; Jupiter's Interior; Jupiter's Magnetic Field and Radiation Belts; Jupiter's Ring System; Jupiter's Satellites; Neptune's Great Dark Spots; Planetary Atmospheres; Planetary Rotation; Planetology: Comparative.

Jupiter's Interior

Categories: Planets and Planetology; The Jovian System

Jupiter, the largest planet in the solar system, is often described as a "gas giant" planet. However, the interior structure of Jupiter is far more complex than a big ball of gas. These interior characteristics are also responsible for many of the planet's observed properties.

OVERVIEW

Jupiter is the largest planet in the solar system, with a mean diameter of about 138,000 kilometers. Jupiter's mass, about 320 times more than Earth's mass. This is greater than that of the rest of the planets combined. However, most of Jupiter's mass is made up of the two lightest elements in the universe: hydrogen and helium. Jupiter is believed to be formed directly from the disk of material swirling together to make the Sun. Thus it is not a surprise to find that Jupiter's composition is very similar to that of the Sun (a star), which is also mostly hydrogen and helium. However, there are significant differences between the structure of Jupiter and that of a star, or even a failed star such as a brown dwarf. The interior structure of Jupiter has not been directly observed. It can only be inferred through mathematical modeling based on observations of the planet. Jupiter does not have a solid surface on which a spacecraft can land, and cloud layers in Jupiter's upper atmosphere shield the interior of the planet from view.

Jupiter is observed to have an oblateness of 0.065 (its equatorial diameter is 6.5 percent greater than the pole-to-pole diameter). This observation constrains the size of any solid core that the planet may have. Mathematical models suggest that Jupiter has a rocky core, with perhaps a nickel-iron inner core, of mass somewhere between 8 and 15 times the mass of the Earth. Due to Jupiter's large mass, though, that is only between 2.5 to 4.7 percent of Jupiter's mass. Despite the core being several times the mass of Earth, the extreme pressure inside Jupiter, nearly 70 million atmospheres, compresses the core to a size on the order of that of Earth. The temperature at the core of Jupiter is believed to be perhaps 22,000 kelvins. It is not known if the core is solid or liquid.

Because Jupiter is in a region of the solar system in which a large number of icy bodies exist, a great number of these bodies must have impacted Jupiter over its history, starting at its formation. These ices, being heavier than the hydrogen and helium that make up the bulk of Jupiter, would have settled toward the deep interior of the planet. Planetary scientists believe that there may exist a layer of this material perhaps 3,000 kilometers thick on top of the core. The ices include frozen ammonia and methane rather than just water ice. The temperature and pressure deep inside Jupiter would ensure that this material is in a liquid state, though, rather than frozen. Therefore, the term "liquid ices" is often used to describe this layer. Great pressure at this depth would make this material behave in ways quite different from the way the same material would behave on Earth.

The bulk of Jupiter is composed of hydrogen and helium. The outermost parts of Jupiter are gaseous, with clouds in the upper 100 kilometers. At great depths inside the planet, the pressure becomes great enough to compress these gases into a liquid state. That pressure is reached at a depth of about 1,000 kilometers below the cloud tops. However, there is no vast ocean of liquid hydrogen under Jovian skies the way that Earth's water collects in oceans. On Jupiter, in fact, there is no clear boundary between the gaseous atmosphere and the liquid interior, because the temperature and pressure inside Jupiter are well in excess of hydrogen's critical point. Beyond the critical point of a substance, there ceases to be a definite phase transition between liquid and gas. Rather, the material takes on a state known as a supercritical fluid. At greater altitudes, the hydrogen in Jupiter is clearly gaseous. At much lower levels, it definitely has more liquid properties, but there is no obvious depth at which the hydrogen becomes liquid. Instead, with increasing depth, the hydrogen becomes more and more like liquid. Though Jupiter is called a "gas giant" planet, the majority of the planet's composition is actually liquid.

At sufficient pressure and temperature,

hydrogen takes on metallic properties. That means that it conducts heat and electricity like any other element on the left-hand column of the periodic table of elements. These conditions are met in Jupiter below a depth of about 7,000 kilometers below the planet's cloud tops. Liquid metallic hydrogen exists from that depth all the way down to the liquid ices at the core. That means that the bulk of Jupiter's mass is in a mantle composed of helium and liquid metallic hydrogen, possibly comprising about two-thirds of the planet.

Jupiter has the strongest magnetic field of any planet in the solar system. At its equator, Jupiter's magnetic field is nearly fourteen times stronger than Earth's magnetic field. Planetary magnetic fields are believed to be created by magnetohydrodynamics in a planet's interior. A dynamo model of planetary magnetic fields shows that a suitable conductor moving in a magnetic field can regenerate that magnetic field, producing a long-lived magnetic field. However, this dynamo effect requires a highly conducting fluid in order to operate. It may be possible for part of Jupiter's core to have a liquid iron region, but that would be not be large enough to account for Jupiter's magnetic field. Rather, Jupiter's magnetic field originates primarily in its liquid metallic hydrogen mantle.

Unlike Earth, Jupiter radiates nearly twice as much energy as it gets from the Sun. This surplus energy is produced by kelvin-Helmholtz contraction. When planets form, a large amount of gravitation energy is released as the materials that form the planet come together. For fluid bodies such as Jupiter, as they radiate thermal energy into space, they contract somewhat. This contraction then compresses the material making up the planet, heating it further. Other gas giant planets besides Jupiter probably also had kelvin-Helmholtz contraction after they formed, but they have long since stabilized at a point where such contraction no longer is a major source of thermal energy. Jupiter is at nearly the perfect mass to extend kelvin-Helmholtz contraction to the longest time possible. If Jupiter had more mass, then it would have compressed faster, until it reached a point of maximum compression. If Jupiter had less mass, like Saturn, then it would have not had sufficient gravity to keep contracting for as long as it has.

KNOWLEDGE GAINED

Much has been learned about Jupiter through observations from Earth. The interior of the planet, of course, cannot be directly measured. Understanding the nature of matter has allowed astrophysicists to make theoretical models of Jupiter's interior; however, it took measurements by spacecraft sent to Jupiter to actually begin to learn more about that planet's interior structure.

Since Jupiter's magnetic field is produced in the planet's mantle, it rotates with the planet's interior. Studies of the magnetic field show that Jupiter's interior rotates once every 9 hours, 55 minutes, and 30 seconds, somewhat more slowly than the rate of rotation of the cloud tops near the planet's equator. Until spacecraft were able to approach Jupiter to measure its magnetic field, astronomers could only guess at its interior rotational period. Jupiter's huge liquid me-

Jupiter Compared with Earth

	Jupiter	*Earth*
Mass (10^{24} kg)	1898.6	5.9742
Volume (10^{10} km^3)	143,128	108.321
Equatorial radius (km)	71,492	6378.1
Ellipticity (oblateness)	0.06487	0.00335
Mean density (kg/m^3)	1,326	5,515
Surface gravity (m/s^2)	23.12	9.78
Surface temperature (Celsius)	–140	–88 to +48
Satellites	63	1
Mean distance from Sun millions of km (mi)	779 (483)	150 (93)
Rotational period (hrs)	9.9250	23.93
Orbital period	11.86 yrs	365.25 days

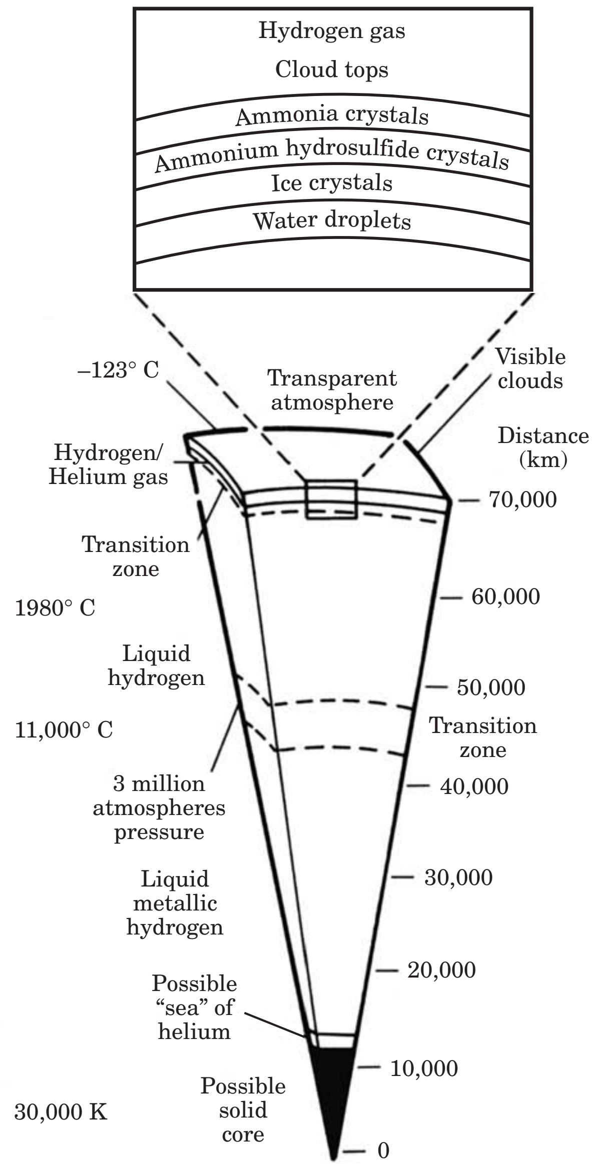
Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.

tallic hydrogen layer produces a magnetic field that is so powerful that Jupiter has a magnetic field stronger than any other planet in the solar system, and Jupiter's magnetosphere is the largest of any of the planets. The existence of Jupiter's powerful magnetic field provided evidence of the metallic nature of hydrogen well before it was produced in the laboratory.

To date, most of the extrasolar planets discovered have been gas giant planets. As the largest gas giant planet in the solar system, studies of Jupiter help to reveal the nature of these planets. Gas giant planets can have masses greater or less than that of Jupiter. However, in size, Jupiter is about as large as a gas giant planet can be. If it had much more mass, gravity would compress it to a smaller volume. Less mass would, of course, make a smaller planet, but its lower gravity would compress the planet less. For example, Saturn has almost 30 percent of Jupiter's mass but more than 80 percent of Jupiter's diameter and more than 55 percent of Jupiter's volume.

Studies of Jupiter's composition have led astronomers to believe that Jupiter may have formed somewhat farther from the Sun than the distance at which it currently orbits. Interactions with other planets, notably Saturn, could have caused Jupiter to migrate inward. This planetary migration also explains observations of extrasolar gas giant planets that appear much closer to their stars than can be explained through current understanding of planetary formation.

Cross-sectional Model of Jupiter



Source: David Morrison and Jane Samz. *Voyage to Jupiter*. NASA SP-439. Washington, D.C.: National Aeronautics and Space Administration, 1980, p. 4.

In late 2008 a research team reported the results of computer simulations based on the properties of hydrogen-helium mixtures exposed to the extreme conditions of temperature and pressure deep inside Jupiter. Their computer simulation also incorporated a core accretion model. In a paper published in the November 20, 2008, issue of *Astrophysical Journal Letters*, these researchers presented an argument that Jupiter's core could be twice as big as previously thought. Their simulation predicted a rocky core perhaps amounting to 5 percent of Jupiter's total mass, making the rocky core equivalent to 14 to 18 Earth masses. That core would have layers of metals, rocky material, methane ice, water ice, and ammonia ice. Like Earth's core, the very center of Jupiter's core would be composed of iron and nickel. This computer simulation could be applied to attempts to understand the cores of the other gas giant planets as well. However, being a computer model, additional observational data and further analysis would be needed before this intriguing claim could achieve complete acceptance from the planetary science community.

CONTEXT

Jupiter and Saturn, the two largest planets in the solar system, probably share a common origin and similar structure. Both formed largely from material that was coming together to form the Sun. Thus, studies of these two worlds allow astronomers to learn more about conditions in the early solar system. Understanding these planets helps astronomers to understand how other planets, including the "rocky" planets such as Earth, form. Jupiter also seems to be similar to exoplanets that have formed around other stars and thus a sort of laboratory for understanding those extrasolar planets.

However, Jupiter is still nearly 600 million kilometers away from Earth, even at its closest approach. Thus, detailed studies have required visits by spacecraft. In total, seven spacecraft have studied Jupiter. Launched in the early 1970's, Pioneer 10 and Pioneer 11, followed in the late 1970's by Voyager 1 and Voyager 2, eventually flew past Jupiter on their way to the outer solar system. Early in 2007, the New Horizons spacecraft flew past Jupiter on its way to Pluto and the Kuiper Belt. The Cassini spacecraft passed by Jupiter December 30, 2000, on its way to Saturn. The Ulysses spacecraft flew by Jupiter in February, 1992, using that planet's gravity to send it into an orbit that permitted it to study the Sun's polar regions. All of these spacecraft studied Jupiter as they went past. The Galileo spacecraft, however, was sent specifically to study Jupiter, orbiting that planet from 1995 to 2003. Galileo also sent an atmospheric probe into Jupiter, the only probe to enter the atmosphere of any of the gas giant planets.

Jupiter is the best studied of the gas giant planets, therefore, but it still holds many mysteries. Its interior must still be investigated through inferences from observations of Jupiter's exterior and of the planet's magnetic field. Debate continues over the exact nature of the planet's interior structure as astronomers pursue additional studies to develop a detailed understanding of this world.

Raymond D. Benge, Jr.

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. New York: Cambridge, 2004. A collection of papers on the current scientific understanding of the planet Jupiter, including a chapter on Jupiter's interior.

Corfield, Richard. *Lives of the Planets: A Natural History of the Solar System*. New York: Basic Books, 2007. This book focuses less on the planets themselves than on the history of planetary exploration and the spacecraft that have studied the planets.

Fischer, Daniel. *Mission Jupiter: The Spectacular Journey of the Galileo Spacecraft*. New York: Copernicus Books, 2001. A good overview of the Galileo mission to Jupiter. Some of the initial findings are given, along with many color photographs of the planet and its moons.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. An excellent college-level introductory astronomy textbook. An entire chapter is devoted to Jupiter and Saturn.

Irwin, Patrick. *Giant Planets of the Solar System: An Introduction*. New York: Springer, 2006. An overview of all four gas giants in the solar system, this text is written at the level of advanced students. It covers all aspects of the planets, including theories of formation, and has a very good bibliography.

McAnally, John W. *Jupiter and How to Observe It*. New York: Springer, 2008. Intended for amateur astronomers, this book focuses on

observations of Jupiter, but it also includes information on the planet itself.

See also: Comet Shoemaker-Levy 9; Europa; Extrasolar Planets; Extraterrestrial Life in the Solar System; Io; Jovian Planets; Jupiter’s Atmosphere; Jupiter’s Great Red Spot; Jupiter’s Magnetic Field and Radiation Belts; Jupiter’s Ring System; Jupiter’s Satellites; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative.

Jupiter’s Magnetic Field and Radiation Belts

Categories: Planets and Planetology; The Jovian System

An understanding of Jupiter’s magnetic field has proved vital to furthering comprehension of this enormous planet’s singular structure, and such knowledge also enriches Earth, planetary, and solar-system science.

OVERVIEW

Since 1610, when Galileo focused his telescope on Jupiter and four of its satellites, this immense planet, orbiting 628 million kilometers from Earth at its closest approach, has received much attention from astronomers. Jupiter’s mean distance from the Sun is 5.2 times the mean distance between the Earth and the Sun, the latter being known as an astronomical unit (AU). After the Sun, Jupiter is the largest object in the solar system, possessing a mass 318 times greater than Earth’s and a diameter 11 times longer. Its volume is thirteen hundred times that of Earth. Because of its prominence, many of Jupiter’s basic characteristics were ascertained centuries ago. Sir Isaac Newton accurately calculated its mass and density, for example. Its radius, diameter, rate of rotation, chemical composition, and singular surface features have similarly been under study for centuries. However, this body, which is neither a star like the Sun nor a terrestrial planet like Earth, has retained certain mysteries, many of which have to do with the composition of Jupiter’s interior and the origins and behavior of its magnetic field and radiation environment.

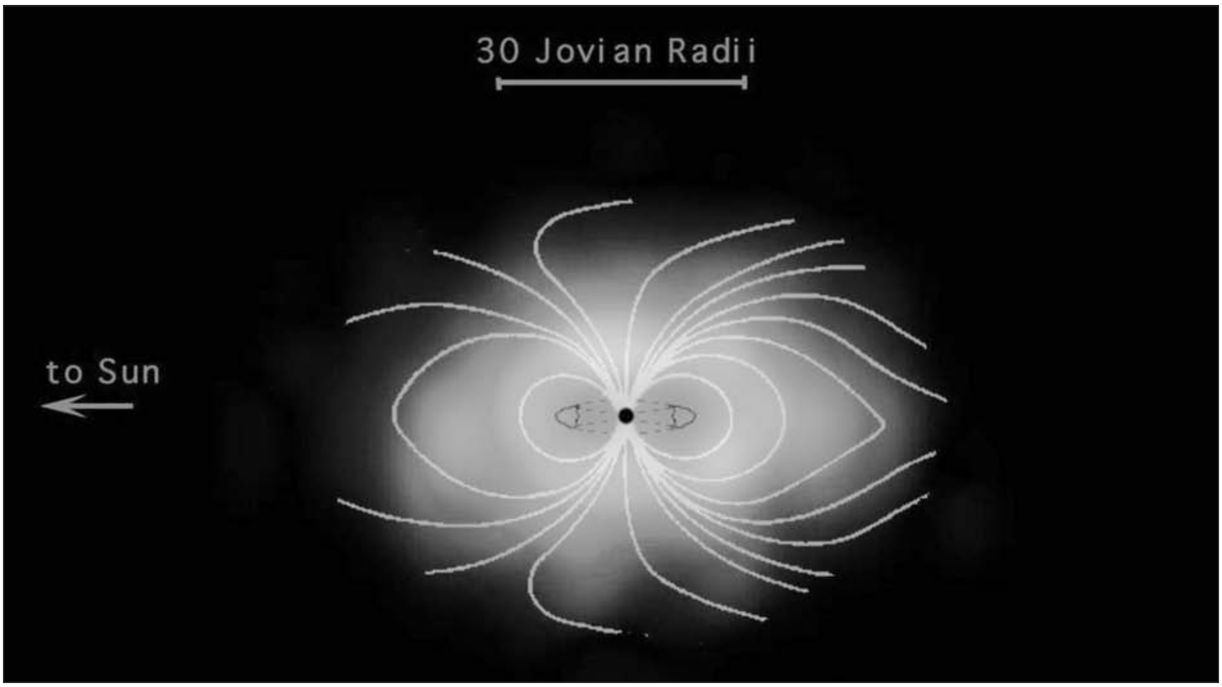
During the 1950’s and 1960’s, prior to investigations by uncrewed spacecraft, radio astronomers gathered approximate data on Jupiter’s magnetosphere, that zone of powerful magnetic influence that surrounds the planet. In 1955, Bernard Burke and Kenneth Franklin, both radio astronomers, found evidence that Jupiter’s magnetosphere was a source of nonthermal radio activity (in contrast to the thermal radiation emitted by all objects with temperatures above absolute zero) at a frequency of 22.2 megahertz (MHz). Other astronomers later noted that this radio activity occurred, if not continuously, at least in a patterned way, at the same point in the planet’s rotation. Such emissions distinguished Jupiter from the other planets and raised provocative questions about the validity of previous theories of radiation.

Several years later, additional radio bursts were picked up by Earth-based radio telescopes in a different portion of the radio spectrum (300 to 3,000 MHz). Unlike the emissions detected earlier, which originated at the planet’s surface, these decametric radiations emanated from within Jupiter’s toroidal region—a region encircling the planet, tilted about 10° from its equatorial belt and extending about 286,400 kilometers from it into space. Within this region and on both sides of the planet at about 140,000 kilometers from its surface there are two “hot spots,” areas of intense radiation activity, which evoked considerable scientific curiosity.

In 1959, two astronomers, Frank Drake and S. Hvatum, identified the source of this toroidal radio activity as synchrotron radiation. Atomic particles within the region were being accelerated to very high speeds by a powerful magnetic field and by changes in the frequency of the electric field. There has been a growing consensus among astronomers that these high-energy particles, electrons that came from the Sun, have been trapped by Jupiter’s external magnetic field, which is 19,000 times stronger than Earth’s magnetic field. Forming radiation belts around the planet, these high-energy particles, moving at high velocities, may produce radio emissions when they strike the top of Jupiter’s atmosphere.

Growing scientific curiosity about these prodigious emissions led to Jupiter’s magnetosphere becoming a subject for investigation during the flybys of Pioneers 10 and 11. Pioneer 10, launched in March, 1972, flew to within 130,000 kilometers of Jupiter on December 3, 1973, securing remarkable photographs of Jovian cloud tops and measuring Jupiter’s radiation belts before sailing farther into space. Pioneer 11, also referred to as Pioneer-Saturn, was launched in April, 1973. On December 2, 1974, it flew to within 42,000 kilometers of Jupiter, photographing the vast but previously unstudied Jovian polar regions. Among other measurements, Pioneer 11 obtained fresh data on Jupiter’s complex magnetic field. The massive flow of data from the Pioneer flights greatly contributed to more precise assessments of the configurations of Jovian radiation belts and the magnetic field, as well as of the extent of the magnetosphere and the distribution of energetic electrons and protons within the planet’s interior. This information, in turn, encouraged fresh ideas about Jupiter’s structure and rotation. In the mid-1970’s, for example, John D. Anderson of the Jet Propulsion Laboratory in Pasadena, California, and William B. Hubbard of the University of Arizona made use of Pioneer data to devise a new model of Jupiter’s internal structure that was consistent with knowledge about both its gravitational field and its magnetic field.

Anderson and Hubbard proposed that beneath the dense and apparently chaotic Jovian atmosphere lies a thick layer of liquid molecular hydrogen, beneath which an even thicker layer of liquid metallic hydrogen exists. The heart of the planet, they proposed, consists of a small, rocky core of iron and silicates heated to temperatures of nearly 30,000 kelvins. Although the presence of such a core could never be proved by gravitational studies, its existence was plausibly deduced from the assumption that if its composition were similar to the Sun’s, it should contain some measure of the same elements. The liquid metallic hydrogen layer presumably extends 46,000 kilometers out from the core, is heated to 11,000 kelvins, and is compressed under a pressure of 3 million Earth atmospheres.

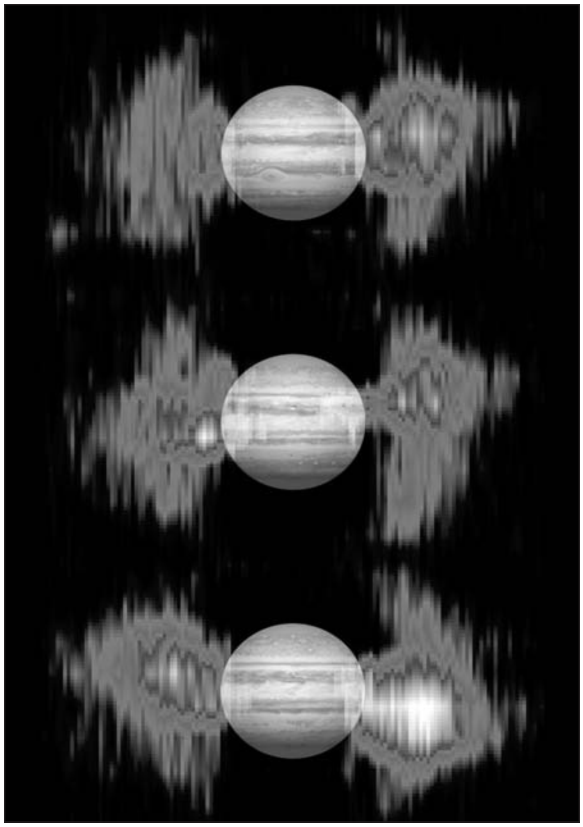


Jupiter’s huge magnetosphere, which covers more space than the Sun, as imaged by instruments aboard the Cassini spacecraft on December 30, 2000. (NASA/JPL/Johns Hopkins University Applied Physics Laboratory)

This layer cannot as yet be experimentally modeled in bulk in a laboratory, yet the construct is plausible: Metallic hydrogen has been created in small amounts in the laboratory. The first to do it were researchers at Lawrence Livermore Laboratory in 1996. In a liquid metallic state, hydrogen molecules break down into atoms and become electrical conductors.

Information from the Pioneer flybys of Jupiter also led to revisions of the hypothesis that the planet's excessive radiation of heat results either from radioactivity or from heat generated by gravitational contraction of the largely gaseous mass of which Jupiter is composed. Since it now appeared that Jupiter was a liquid body (and liquids are all but incompressible), it seemed likely that its excess radiated heat was merely a residue of the heat generated when the planet coalesced from the solar nebula. The implication would be that the planet's original thermal energy is continuously finding its way to the surface and, in the process, creating convection currents, the rising of hot gases or liquids and the downward movement of cold liquids or gases in the planet's interior. Such grand-scale convection currents, as described by John H. Wolfe, who served as chief scientist for the Pioneer missions to Jupiter, could constitute a mechanism for generating the Jovian magnetic field. As primordial heat rose through Jupiter's liquid metallic hydrogen, stirring it, the Coriolis force affected the resulting convection currents. The Coriolis force arises from planetary rotation and deflects other forces in motion—depending on whether they are north or south of the body's equator—either to the right or to the left, much as a person walking across a moving carousel appears thrown off course as seen in an inertial reference frame. Deflected convection currents in such circumstances would set up loops of electric current, which could and may create a magnetic field.

The magnetosphere of Jupiter was determined to expand and contract under pressure from the solar wind. Where an equilibrium existed between magnetic forces and the solar wind's incident stream of charged particles, a planetary magnetopause developed. Pioneer data confirmed the magnetopause to be as far out from the Jovian atmosphere as 7,135,000 kilometers and as close to its atmospheric layer as 3,565,000 kilometers. Just as there is a bow shock wave, there is also a magnetotail, which, much like a ship's wake, extends several hundred kilometers behind Jupiter in a direction away from the Sun.



Mapped by the Cassini spacecraft, Jupiter's inner radiation belts, showing three views during a ten-hour rotation of Jupiter. The image of the planet has been superimposed. Because Jupiter's magnetic field is tilted in relation to the planet's poles, the radiation belts wobble during the planet's rotation. (NASA/JPL)

Like Earth's magnetic field, the Jovian field is dipolar, but its magnetic axis is tilted between 9.5° and 10.8° from the planet's rotational axis, a displacement of about 7,000 kilometers from Jupiter's center. The strength of the magnetic field, as measured at Jupiter's cloud tops, varies from 3 to 14 gauss (a unit of magnetic field strength), extremely powerful by comparison with the 0.3- to 0.8-gauss strength of Earth's magnetic field at the surface. Probably because of the still unknown circulation patterns of the liquid metallic hydrogen in the Jovian interior, Jupiter's magnetic field is far more complex than Earth's. In addition to its dipolar field, the Jovian field, according to Pioneer 11 data, also has quadrupole and octupole movements. That is, components of the main field have four and eight poles, respectively, although these are much weaker than the main dipolar field and have been detected only close to the planet.

Further complications arise from the motion of at least five Jovian satellites that orbit within its magnetosphere. In the course of their orbits, Io, Ganymede, Europa, Callisto, and Amalthea absorb highly charged particles that otherwise might be trapped by Jupiter's magnetosphere. Thus these satellites clear channels through Jupiter's radiation belts. Io sputters material into space that forms a torus of charged particles along its orbit about Jupiter. This torus consists of sulfur, sodium, oxygen, and a few lesser constituents ejected from volcanic activity. Ganymede complicates Jupiter's magnetosphere since that satellite itself possesses a magnetic field and therefore has a magnetosphere within Jupiter's own magnetosphere.

Perhaps because Io in particular possesses an ionosphere that provides it with a conductive fluid, not only has it trapped charged particles otherwise destined for Jupiter's magnetosphere, but it also produces and accelerates charged particles. When Io is in a fixed position along an Earth-Jupiter line, radio emissions increase; thus, it is believed that its charged particles are an additional source of these emissions. Amalthea has also shown peculiarities as it orbits through the Jovian magnetosphere. Charged particles in its magnetic field unexpectedly do not increase in density toward Jupiter's magnetic equator or to its "surface." Instead, particle density varies widely at many different points.

One of the major reasons that the Galileo spacecraft was outfitted with a sophisticated magnetometer was to attempt to determine how plasma is transported through Jupiter's magnetosphere. Naturally the magnetometer was also intended to provide precise determinations of spatial and temporal variations of Jupiter's magnetic field, and the extent and shape of the planet's complex magnetosphere as well as its interactions with the solar wind.

KNOWLEDGE GAINED

Between the recording of Jovian radio emissions in the 1950's and the Pioneer and Voyager robotic space flights more than twenty years later, knowledge of the Jovian magnetic field increased by quantum leaps. New data not only have confirmed or disproved older theories about the field's origins, extent, and inner and outer complexities but also have led to more consistent and plausible theories about the structure of Jupiter itself—indeed, of other planets as well.

Sources and at least some causes of Jupiter's copious radio emissions have been identified. Based on hard, if still incomplete, evidence gathered by American spacecraft, the configurations of the magnetosphere, magnetopause, magnetosheath, and magnetotail have been delineated. The magnetic field's axis and its location have been defined. Quadrupole and octupole fields within the overall dipolar field have been discovered. The strength of the field and fluctuations within it have also been measured. Investigations of the magnetosphere have led to new theories concerning the makeup of the Jovian interior and some of the convective and conductive functions of the liquid metallic hydrogen that composes much of it.

Further, many of the special characteristics of the outer magnetosphere have been explained as the result of two principal mechanisms. First, as observed by James Van Allen (discoverer of Earth's Van Allen radiation belts), a mass of low-energy plasma trapped in the Jovian magnetic field has created pressures that have inflated the field as if it were a balloon. Second, because of interactions with the planet's rotating magnetic field, the plasma corotates (over a period of 9 hours, 55.5 minutes), creating a centrifugal force that contributes to the outward pressure. Plasma analyzers on the two Voyager spacecraft indicated that the plasma originated from gases—principally sulfur dioxide and hydrogen sulfide with sodium and oxygen in lesser amounts—vented by Io's vigorous volcanic activities. This plasma is responsible for the Io torus, a phenomenon unique to Jupiter's magnetosphere. That is, Io is surrounded by a doughnut-shaped band of excited charged particles, some captured from the solar wind as it orbits Jupiter and some produced and agitated by Io's own environment.

It is also understood now that Io and the other inner satellites, in the course of their orbits, attract many particles that enter Jupiter's inner magnetosphere and thereby limit that region's population of trapped particles. Particles trapped in the Jovian magnetosphere also affect the chemistry of the planet's environment. Andrew Ingersoll has shown that trapped particles rain down from the magnetosphere into the Jovian atmosphere. There lightning and other charged particles break down dominant chemical species, thus keeping a balance between production and breakdown of hydrocarbons in the Jovian atmosphere.

The Galileo spacecraft provided long-term measurements of Jupiter's magnetic field and the interaction of the larger satellites with that complex structure. Magnetic measurements combined with gravitational information refined the understanding of Jupiter's interior structure.

CONTEXT

Jupiter attracts scientific attention because it is big, stunningly dynamic in visual wavelengths, and noisy in the radio range. By far the largest body in the solar system after the Sun, it is a natural target of interest for the astronomer's gaze and telescopic observation. Consequently, the planet's general outward appearance—its cloud cover, its bands, and its fascinating Great Red Spot—have been closely observed over several centuries.

An emitter of abundant radio emissions, Jupiter was sufficiently noisy to arouse the curiosity of radio astronomers, who were then led to explore the spectrum of these emissions and theorize about the planet's structure. By the 1950's, thanks to analysis in the visible, infrared, and radio frequency portions of the electromagnetic spectrum, much was known of Jupiter's chemical composition. The light gases hydrogen and helium were dominant, and the sheer size of Jupiter meant that it would require billions of years to dissipate them. Some believed that beneath the planet's dense cloud cover there were mountains, while others believed the planet to be entirely gaseous. In the light of the vast quantities of information gained by the early 1980's, much of the pre-1950 understanding of the planet seems rudimentary, even laughable.

Between the 1950's and the early 1980's, several broad intellectual currents inspired scientific interest in planetary science, in Jupiter, and thereby ultimately in the planet's magnetosphere. One such current was general scientific acceptance of a theory of the solar system's origin that postulates condensation from a disk-shaped nebula of gas and dust. This theory proposes that about 4.6 billion years ago, a cloud of interstellar gas or dust, overwhelmed—as Carl Sagan and others have suggested—by an exploding star, collapsed and condensed to form the solar system. The central mass in the interstellar formation, contracting under its own gravity, produced such prodigious heat that it generated a thermonuclear reaction, from which the Sun evolved. Lesser masses—Earth, for example—experienced less heating and so, as planets, moons, asteroids, and comets, bathed in and reflected the Sun's light. Substantiation of this theory's accuracy, validity, and reliability required a fresh empirical examination of each major object in the solar system.

With the dawn of the space age, it became possible to make in situ observations or at least the nearest equivalent. Chronologically speaking, data collected by the flybys of Pioneers 10 and 11 and Voyagers 1 and 2, by Galileo in orbit about Jupiter, and by Cassini and New Horizons passing through the Jupiter system, en route to other destinations, vastly increased the storehouse of knowledge concerning Jupiter's magnetosphere. Contributions of these exploratory missions to astronomy and comparative planetology helped advance a greater understanding of Earth's relationship to the rest of the solar family.

Clifton K. Yearley

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. Cambridge, England: Cambridge University Press, 2007. A comprehensive work about the biggest planet in the solar system, comprising a series of articles by experts. An excellent repository of photography, diagrams, and figures about the Jupiter system and the various spacecraft missions that unveiled its secrets.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. Filled with color diagrams and photographs, this is a popular work on solar-system astronomy and planetary exploration through the Mars Pathfinder and Galileo missions. Accessible to the astronomy enthusiast, it provokes excitement in the general reader and communicates an appreciation for the need to understand the universe around us.

Cole, Michael D. *Galileo Spacecraft: Mission to Jupiter*. New York: Enslow, 1999. Provides a full description of the Galileo spacecraft, its mission objectives, and science returns through the primary mission. Particularly good at describing mission objectives and goals. Suitable for a younger audience.

Fimmel, Richard O., James Van Allen, and Eric Burgess. *Pioneer: First to Jupiter, Saturn, and Beyond*. NASA SP-446. Washington, D.C.: Government Printing Office, 1980. A classic, informative, and readable account of the first spaceflight into the Jovian environment, offering an excellent synthesis of its amazing findings. Includes illustrations, tables, a modest bibliography, and an index.

Harland, David H. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. New York: Springer Praxis, 2000. This book provides, in a single work, virtually all of NASA's press releases and science updates during the first five years of the Galileo mission. Contains an enormous number of diagrams, tables, lists, and photographs. Provides a preview of the Cassini mission (which did not end until after publication).

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Jupiter is discussed in comparison to the other planets, as the text takes a comparative planetology approach rather than providing individual chapters on each planet in the solar system.

Ingersoll, Andrew P. "Jupiter and Saturn." In *The Planets*, edited by Bruce Murray. San Francisco: W. H. Freeman, 1983. An authoritative discussion that is profitable for both specialists and educated laypersons. Includes superb photographs and graphics. At the end of the book, there are biographies of Ingersoll and other authors, as well as a select bibliography for this and other chapters. The volume covers only what was being learned from the Pioneer and Voyager spacecraft on their flybys of Jupiter, but it is an excellent source for early understandings of the two gas giants.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Accessible to the serious general audience.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar-system astronomy suitable for an introductory college course but accessible to nonspecialists as well. Filled with supplemental learning aids and solved student exercises. A companion Web site is available for educator support.

Stone, Edward C., and A. L. Lane. "Voyager 2 Encounters with the Jovian System." *Science* 206 (November 23, 1979): 925-927. Technical, but readily understandable by readers with a basic background in mathematics and physics. Discusses data generated about the Jovian magnetosphere and other features of the planet (and its moons). Stone was chief scientist for the Voyager mission. Contains illustrations, tables, and a brief bibliography.

See also: Auroras; Brown Dwarfs; Comet Shoemaker-Levy 9; Europa; Io; Jovian Planets; Jupiter's Atmosphere; Jupiter's Great Red Spot; Jupiter's Interior; Jupiter's Ring System; Jupiter's

Satellites; Kuiper Belt; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Saturn's Atmosphere.

Jupiter's Ring System

Categories: Planets and Planetology; The Jovian System

Jupiter's ring system consists of four relatively dull, ethereal rings composed of submicron- to micron-sized dust grains. This system provides important clues and insights into the processes that are involved in the generation of circumstellar disks around planets. In the case of Jupiter, the primary mechanism that produces and replenishes its rings is dust generated when interplanetary meteoroids collide with four of Jupiter's small, inner satellites.

OVERVIEW

Trailblazing missions to explore Jupiter and Saturn were conducted by Pioneer 10 and 11 and Voyager 1 and 2. When Pioneer 11 flew by Jupiter in 1973-1974, observations of rather rapid variations in the number of charged particles orbiting Jupiter at specified distances from the planet suggested the possibility of a ring system that might be absorbing the particles. Although the Pioneer spacecraft were not sufficiently stabilized to facilitate taking images, Voyager 1 and 2 were. On March 4, 1979, an overexposed image from Voyager 1 finally confirmed the existence of a ring system around Jupiter, a result long anticipated by astronomers. Voyager 2 cameras captured numerous pictures of Jupiter's ring system at geometries and resolutions that were previously unobtainable. Three separate rings were discovered, the central Main Ring, the inner Halo Ring, and the outer Gossamer Ring. These rings exist within the Roche limit, the distance from the planet to where tidal forces prevent ring particles from forming into aggregates due to gravitational attraction.

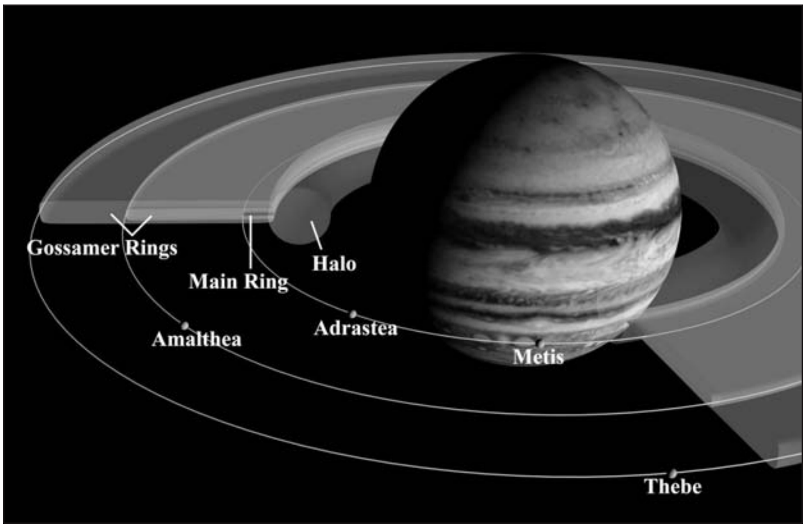
On October 18, 1989, the 2.7-ton Galileo spacecraft, consisting of the main body orbiter and a probe, was launched. On December 7, 1995, the orbiter reached Jupiter and made thirty-four orbits around the planet before plunging into the planet's atmosphere in 2003. Using a solid-state imaging camera, high-quality images were taken of Jupiter's satellite-ring system. After careful analysis of the pictures, it was concluded that Jupiter's ring system is formed from dust generated as high-speed interplanetary micrometeoroids collide with the planet's four small inner satellites (sometimes called moonlets)—Metis, Adrastea, Amalthea, and Thebe—which orbit within the rings. A totally unexpected result was that the outermost Gossamer Ring consisted of two rings, one embedded inside the other. The outermost Gossamer Ring was bounded by the orbit of Thebe and named the Thebe Gossamer Ring, while the innermost one was bounded by the orbit of Amalthea and named the Amalthea Gossamer Ring.

The Cassini spacecraft, designed to obtain high-resolution images of planetary ring systems, made its closest approach to Jupiter on December 30, 2000. It imaged Jupiter's rings using different wavelengths that provided further constraints on the size, distribution, shapes, and composition of the particles within the rings. The reddish colors of the Jovian ring particles indicated a silicate or carbonaceous composition, just like that of the small embedded satellites. Images showed that the particles in Jupiter's rings were nonspherical. Cassini images also captured the motion of the two Gossamer Ring satellites, Thebe and Amalthea.

Images of Jupiter's rings taken by the New Horizons spacecraft in early 2007 confirmed that the dusty Jovian ring system was being replenished continually from embedded source bodies. The Main Ring was found to consist of three ringlets, one just outside the orbit of Adrastea, one just inside the orbit of Adrastea, and one just outside the orbit of Metis. Boulder-sized clumps, consisting of a close-paired clump and a cluster of three to five clumps, were discovered in the Main Ring just inside the orbit of Adrastea. Although the origin and nonrandom distribution of the clumps remain unexplained, they are confined to a narrow belt of motion by the gravitational influence of the two innermost satellites of Jupiter. New Horizons images established a lower limit to the diameter of Jupiter's moons of 0.5 kilometer.

From observations, measurements, and numerical modeling methods applied to data collected from Voyager 2, Galileo, Cassini, New Horizons, the Hubble Space Telescope (HST), and the Keck telescope, it has been concluded that Jupiter's rings are extremely tenuous and contain significant amounts of short-lived dust. In addition to the gravitational perturbations produced by the small satellites embedded within and bounding Jupiter's ring system, the dynamics of its faint, ethereal dusty rings are dominated by effects that involve electromagnetic forces, solar radiation pressure, and various drag forces. In the process of conserving angular momentum, the rapid spin rates tend to flatten the rings. The relatively bright, narrow Main Ring has a rather sharp outer boundary that coincides with the orbit of Adrastea. Just inside this boundary is the orbit of Metis. Since the Main Ring extends only inward from these small source moonlets, it has been concluded that particles in the Main Ring must drift inward. The width of the Main Ring is approximately 6,440 kilometers, with a thickness that varies between 30 and 300 kilometers. Dust size ranges from 0.5 to 2.5 microns in diameter.

Interior to the Main Ring lies a thick torus of particles known as the Halo Ring. Its thickness is primarily determined by Jupiter's very strong magnetic field operating on the ring's submicron dusty grains. The thickness of the Halo Ring is approximately 12,500 kilometers, while its width is about 30,500 kilometers. In visible light, the Halo Ring has a bluish color. The very faint Almathea Gossamer Ring has an estimated width of 53,000 kilometers and a thickness of 2,500 kilometers. It has been imaged from the Earth using the Keck telescope. It appears brighter near its top and bottom edges and also brightens toward Jupiter. The dust grain size in this ring is similar to that in the Main Ring. The faintest Jovian ring, the Thebe Gossamer Ring, has a width of approximately 97,000 kilometers and a thickness of 8,500 kilometers. Dust grain size varies from 0.2 to 3.0 microns. The Thebe Gossamer Ring is observed to extend beyond Thebe, which is apparently due to coupled oscillations produced by time-varying electromagnetic forces that cause the ring to extend outward. The thickness of each Jovian ring is primarily controlled by the inclination of the orbit of its embedded moonlet.

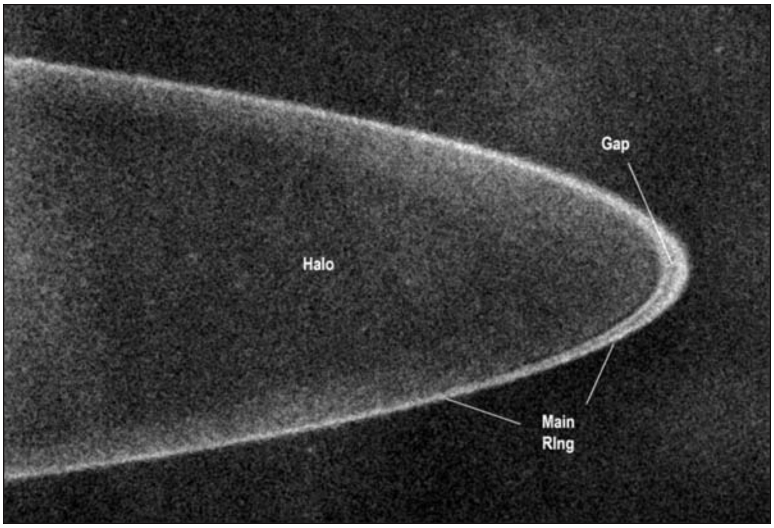


The diagram displays the main components of Jupiter's ring system as well as the small inner moons, from which the dust of the rings originates. (NASA/JPL/Cornell University)

KNOWLEDGE GAINED

The existence of Jupiter's ring system was unambiguously determined in March, 1979, by Voyager 1. Until that time, most astronomers and astrophysicists were confounded as to why Saturn had a ring system but Jupiter did not. In July, 1979, more detailed images from Voyager 2 showed that the dull, diffuse ring system of Jupiter consisted of three separate rings. The ring system exists within an intense radiation belt of electrons and ions that are trapped in Jupiter's magnetic field. Resulting drag forces play an important role in determining the motion of the ring particles.

Images obtained from the Galileo spacecraft between 1995 and 2003 provided increasing detail about Jupiter's rings. The shape, width, thickness, optical depth, and brightness of each ring were determined, as well as dust spatial densities, grain sizes, and grain collision speeds. Jupiter's faint, dark, narrow rings (albedo about 0.05) consist of submicron- to micron-sized rock fragments and dust but do not contain ice, as do Saturn's rings. The number of separate rings in Jupiter's ring system was found to be four when it was determined that the Gossamer Ring consisted of two distinct rings. Further constraints on the composition, distribution, size, and shape of particles within Jupiter's rings were established in 2000 by the Cassini probe. In 2007, images from the New Horizons spacecraft revealed the fine structure of Jupiter's Main Ring. It consists of three ringlets and contains two families of boulder-sized clumps.



Jupiter's ring system was imaged by the New Horizons Long Range Reconnaissance Imager on February 24, 2007. (NASA/Johns Hopkins University Applied Physics Laboratory/Southwest Research Institute)

From the variety of measurements, observations, and analysis of collected data, Jupiter's ring system is now the best understood prototype of planetary ring systems that consist of thin, diffuse sheets of dusty debris that are primarily generated by small source moonlets. The relative motion of the dust grains within Jupiter's rings and the orientation of the orbits of the rings are primarily controlled by three processes: the spinning, asymmetric, very strong magnetic field of Jupiter; the absorption, reemission, and scattering of solar radiation energy by the dust particles, which produces momentum changes that induce orbital changes; and drag forces on the grains produced by solar radiated photons, as well as by ions and atoms that are orbiting around Jupiter. Since dust particles are continually being removed from Jupiter's rings by these processes and then replenished by dust from the four inner satellites, the dust grains existing in the rings are estimated to be relatively young, probably much less than one million years old. As dust particles are ejected from the moonlets, the particles enter orbits like those of the moonlets, which causes the rings to wobble up and down as they orbit around Jupiter's equator. Micrometeoritic impacts that generate dust from the moonlets also color, chip, erode, and fragment the dust particles within the rings.

CONTEXT

Spacecraft flybys and orbiters of Jupiter and Saturn have greatly increased the scientific understanding of planetary rings. Numerical methods have been employed to simulate the physical processes occurring within Jupiter's rings by including collisional, gravitational, and electromagnetic interactions among the orbiting ring particles. Resulting models are providing keys to help guide observational strategies for future space missions.

The ring system of Jupiter provides insights into the characteristics of flattened systems of gas and colliding dust particles that are analogous to those that have eventually resulted in the formation of solar systems. In particular, Jupiter's rings offer an accessible laboratory for observing, measuring, and modeling the ongoing processes similar to those associated with the circumstellar disks that were most likely active in the solar nebula disk when the solar system containing the Earth was formed.

Further analysis, detailed examination, and numerical modeling of the data acquired by the Cassini probe and New Horizons spacecraft should provide more high-resolution maps, identify the detailed radial structure of Jupiter's ring system, and reveal invaluable time-variable features associated with the evolution of the rings. Future observations and measurements will offer new insights into the dynamic forces that shape and maintain these fascinating structures. The National Aeronautics and Space Administration (NASA) is considering a mission to Jupiter to explore the planet and its satellite-ring system in detail from a polar orbit. NASA also plans to develop small spacecraft with the capability of hovering over the rings of Jupiter and Saturn, which should provide the necessary additional data and insights for producing refined models and an advanced understanding of planetary ring structures and why they vary vastly among the gas giants.

Alvin K. Benson

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. Cambridge, England: Cambridge University Press, 2004. The physical characteristics, temperature, atmospheric makeup, and satellite and ring systems of Jupiter are clearly analyzed. A description of spacecraft missions to Jupiter, numerous high-quality, full-color photographs, and many figures, tables, and diagrams elucidate a tour of the Jovian system.

Elkins-Tanton, Linda T. *Jupiter and Saturn*. New York: Chelsea House, 2006. Discusses the role that Jupiter has played in advancing our understanding of the planets that orbit the Sun. Clearly describes the discovery of Jupiter's ring system, details the Jupiter satellite-ring system, explains why the rings exist, and contrasts the ring systems of Jupiter and Saturn. An appendix lists all the known satellites of Jupiter and Saturn.

Esposito, Larry. *Planetary Rings*. Cambridge, England: Cambridge University Press, 2006. This treatise covers all aspects of planetary ring systems, including ring history, physical processes involved in ring evolution, and mathematical models used to describe them. In particular, Esposito discusses Jupiter's ring-satellite system, the age of Jupiter's rings, and the size distribution of Jupiter's rings. The text is clearly written, illustrated with many diagrams and images, and geared for a wide reading audience.

Harland, David M. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. Chichester, England: Springer Praxis, 2000. A detailed account of the long trek to Jupiter by Galileo and its five years of exploration within the Jovian system. Spectacular results are presented from the observations and measurements of the satellite-ring system of Jupiter. Written for general readers, undergraduates, and faculty alike, the book is very well illustrated and includes references to many relevant Web sites.

Krüger, Harald. *Jupiter's Dust Disc: An Astrophysical Laboratory*. Aachen, Germany: Shaker-Verlag, 2003. On the basis of spacecraft observations and measurements made by Galileo, Krüger delineates the physical processes and mechanisms involved in producing Jupiter's ring system. The data clearly indicate that the Gossamer Ring material comes from Jupiter's small moonlets Thebe and Amalthea.

Miner, Ellis D., Randii R. Wessen, and Jeffrey N. Cuzzi. *Planetary Ring Systems*. New York: Springer Praxis, 2006. Provides comprehensive coverage of the scientific significance of ring systems, ring characterization and comparison, and the history of the discovery of planetary ring systems, including observations of Jupiter's ring system from Voyager 1, Voyager 2, and Galileo. Various theories for the formation of planetary ring systems are explored.

See also: Europa; Io; Jovian Planets; Jupiter's Atmosphere; Jupiter's Great Red Spot; Jupiter's Interior; Jupiter's Magnetic Field and Radiation Belts; Jupiter's Satellites; Planetary Ring Systems; Planetary Satellites; Saturn's Atmosphere; Saturn's Ring System.

Jupiter's Satellites

Categories: Natural Planetary Satellites; Planets and Planetology; The Jovian System

Jupiter has many natural satellites, among which are some of the largest and most intriguing of the solar system. Pioneer 10, Pioneer 11, Voyager 1, Voyager 2, and Galileo collected photographs and other vital information about several of these Jovian satellites, which have been used to describe the histories of these bodies and the solar system in general.

OVERVIEW

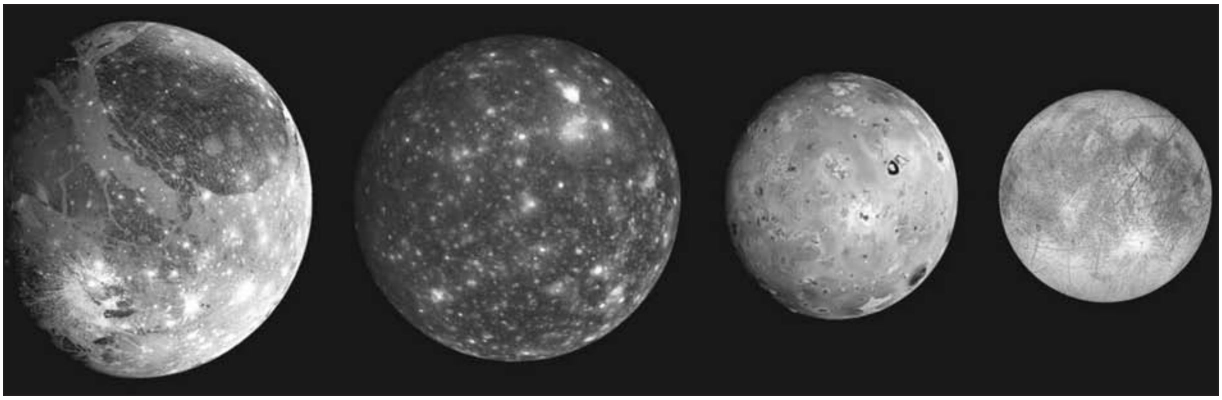
The four largest satellites of Jupiter, Io, Callisto, Europa, and Ganymede, were discovered in 1610 by Galileo Galilei and have been objects of curiosity for astronomers since that time. By 1990 astronomers had discovered sixteen moons around Jupiter—fourteen using Earth-based telescopes and two using photographs from Voyager. One of the fourteen was confirmed by the Voyager mission. After the arrival of the Galileo spacecraft in orbit about Jupiter, and in cooperation with space-based and Earth-based telescope observations, the discovery of Jovian satellites increased regularly. As of 2008, Jupiter had sixty-three recognized satellites. Satellites of the outer planets are classified as "regular" if their orbits are nearly circular and are in, or near, the plane of the equator of the planet. "Irregular" satellites have highly elliptical orbits, have the orbital plane tilted relative to the equator, or revolve in a retrograde (westward) direction.

Jupiter's satellites can be divided into three groups. The largest satellites, the often-called Galilean moons, are all regular satellites, as are the four small satellites whose orbits are nearer to Jupiter than is Io's. Eight irregular satellites lie beyond Callisto, and four of them have retrograde revolutions. Only the Galilean satellites and Amalthea have been examined closely by spacecraft. The other satellites are too small and too distant from the probes that have passed through the Jovian system to be effectively studied.

Many of the outer planets are known to have icy surfaces rather than the rocky surfaces that charaacterize the terrestrial planets (Mercury, Venus, Earth, and Mars) and their satellites. Ice behaves somewhat like a plastic unless its temperature is less than 133 kelvins—that is, it will flow and not retain a structure over long periods of time. This process is called plastic deformation. Rocky surfaces do not undergo plastic deformation. It is expected, therefore, that the terrestrial planets and satellites would retain structural features such as craters indefinitely, while an icy satellite might not be able to do the same unless the temperature is always below 133 kelvins.

Amalthea is the innermost of Jupiter's sizable satellites. It is small and elongated, with a length of 270 kilometers and an average diameter of 155 kilometers. Its long axis is pointed toward Jupiter throughout its orbit. The composition of its dark red surface and of its bulk is unknown. Four features were found by Voyager—two craters, Pan and Gaea, and two mountains, Mons Ida and Mons Lyctas. Nothing definite is known about these features. The scarred surface reflects a history of bombardment.

Io is the Galilean satellite nearest Jupiter. It is a rocky satellite with little or no remaining water. In fact, it may be the driest place in the outer solar system. Its density is similar to that of Earth's moon, so it is assumed that its bulk composition must also resemble that of the Moon. Its surface features, however, are dramatically different from those of the Moon. The surface displays materials that are yellow, red, brown, black, and white. Io shows no cratering but contains layered lava plains, volcanic mountains, calderas, and other evidence of volcanism. Of the nine active volcanoes found by Voyager 1, seven were still active when Voyager 2 passed. The highest mountains were about 9 kilometers tall, making them the largest found on the Jovian satellites. The polar areas are darker than the remainder of Io and show evidence of long periods of deposition, faulting, and erosion. Significant changes in Io's volcanic activity were noted when the Galileo spacecraft arrived in 1996, fifteen years after Voyager 1 passed by.



A composite image showing the "Galilean" satellites of Jupiter—the four largest (left to right): Ganymede, Callisto, Io, and Europa. The Callisto image is from Voyager's 1979 flyby of Jupiter; the rest were taken in 1996 by the Galileo spacecraft. (NASA/JPL/DLR)

Io's volcanic activity takes on several forms. Plume activity is associated with the region 45° north and south of the equator. Material in the plumes is ejected at a velocity of 3,200 kilometers per hour. Much of this material is sulfur dioxide, which then crystallizes and falls back to the surface as a white solid. The material in the plumes is apparently carried to the surface along fissures in the surface rather than through pipe vents like those characteristic of volcanoes found on Earth. Some of Io's volcanoes produce a lava flow rather than a plume. Hot spots exist on the surface where a large amount of energy is being transferred from the interior of Io but where the quantity of volatile material available is insufficient to produce a plume. The location of such hot spots is not limited to the band in which the plume volcanoes are found. As a consequence of the extensive volcanic activity, the surface of Io changes rapidly. Significant details changed between the flyby of Voyager 1, the flyby of Voyager 2, and the arrival of the orbiting Galileo spacecraft. Galileo discovered that Io has an iron core.

Europa also provided surprises for the Voyager scientists. Density measurements imply that the moon is about 90 percent rocky core and 10 percent water ice as the crust. Because of its nearness to Jupiter, it was expected to have an extensively cratered surface. Voyager photographs of Europa, however, show one of the smoothest surfaces in the solar system. Very few craters are visible. The surface is marked with long, narrow lines that resemble the cracking pattern of an egg, along with dark spots and mottling. Lines are regions where the crust cracked and water from the interior was squeezed out onto the surface. A reflectivity of 64 percent is an indicator of quite pure water ice on the surface of Europa. Since an old surface would be expected to collect dust and other space debris and produce a dirty surface with low reflectivity, the pure water ice implies that there is some process taking place that continually renews Europa's surface. After Galileo's close flybys in 1996 and 1997, scientists concluded the ice crust is no thicker than 150 kilometers and floats on a liquid ocean.

Ganymede is the largest of the Jovian satellites, although it is not much larger than Callisto. Ganymede has a metallic core surrounded by a layer of ice and silicates, and its crust is probably thick water ice. The metallic core generates a magnetic field. About one-half of Ganymede's surface has a dark, cratered terrain, while the remaining half is much lighter and has fewer craters.

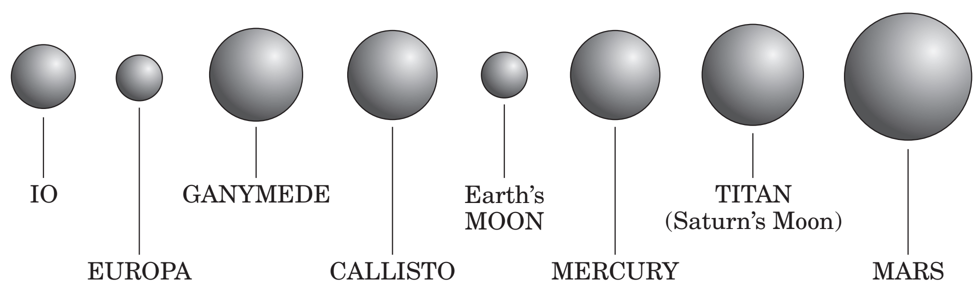
Lighter areas of Ganymede have a series of parallel mountains and valleys reminiscent of the low Appalachian Mountains in North America. These may have developed from long cracks or faults that are separated by strips of land that have alternately been lifted or depressed. Depressed areas could have flooded with liquid water from the interior. This type of mountain building results from tension in the crust rather than the compression of the crust that is believed to have caused the Appalachian Mountains to be uplifted. Reflectivity of the lighter areas is about 40 percent, while in the more heavily cratered areas it is only 25 percent. Bright ray craters are found in all parts of the moon. (A ray crater is one in which the debris cast out of the crater at the time of impact is distributed along radial directions like rays or beams.) These craters are evidently the result of water from under the crust that has splashed out on the surface when some object collided with Ganymede. A dome 260 kilometers in diameter and 2.5 kilometers in height which is surrounded by a number of small craters may be evidence of water volcanism. Impacts that formed the small craters may have weakened the crust, so that water flowed up through newly formed cracks and holes. There are very few large craters, but some ghost craters exist. These show the details of the crater, but actual physical features such as walls have disappeared because of plastic deformation of Ganymede's surface.

Callisto is the outermost of the Galilean satellites and probably the least active. Its surface temperature ranges from 150 kelvins at noon to 100 kelvins before dawn. At temperatures as cold as these, a layer of ice only 1 meter thick would take 4.5 billion years to evaporate. The dark, heavily cratered surface has a reflectivity of only 18 percent, which implies that the surface ice has a low purity. Callisto's landscape is almost exclusively the product of impacts. The crater features are subdued, and no large impact basins exist. This relative smoothness is probably the result of the plastic deformation of the crust. In the Valhalla basin, visible rings on the surface may be the remnant of a very large impact basin. Few ray craters are visible.

KNOWLEDGE GAINED

The composition of the surface of Io has been reasonably well established. The white solid is sulfur dioxide condensed from plume activity. The remainder of the surface is largely elemental sulfur forms, which can have a variety of colors, depending upon the process occurring during solidification. Sodium has been detected in the space around the satellite. The existence of active volcanoes on Io requires the presence of molten material under the surface. Tidal action of Jupiter and of other Jovian satellites upon Io is probably the main source of heat energy that generates molten material. Although Jupiter exerts a very large attractive force on Io, the gravitational attractions of the other satellites cause the surface of Io to be pulled in competing directions. Friction resulting from the flexing of the surface causes the material beneath the crust to melt. The thickness of the crust is not precisely known, but it is thought to be minimal. A sulfur sea may exist below the crust. An additional source of heat is electric current induced in the iron sulfide core of the satellite as it interacts with the strong magnetic field of Jupiter. A portion of the radio waves emitted by the Jovian system originate from Io's interaction with the

Sizes of the Jovian Moons Relative to Other Bodies



Some Facts About the Major Jovian Satellites

	Mass (10^{20} kg)	*Radius (km)*	*Mean Density (kg/m^3)*	*Orbital Period (days)*
JUPITER'S MAJOR MOONS				
Io	893.3	1,821.3	3,530	1.77
Europa	479.7	1,565	2,990	3.55
Ganymede	1,482	2,634	1,940	7.15
Callisto	1,076	2,403	1,851	16.69
SATURN'S MAJOR MOONS				
Mimas	0.375	209	1,140	0.94
Enceladus	0.73	256 × 247 × 245	1,120	1.37
Tethys	6.22	536 × 528 × 526	1,000	1.89
Dione	11.0	560	1,440	2.74
Rhea	23.1	764	1,240	4.52
Titan	1,345.5	2,575	1,881	15.94
Hyperion	0.2	185 × 140 × 113	—	21.28
Iapetus	15.9	718	1,020	79.33
URANUS'S MAJOR MOONS				
Miranda	0.66	240 × 234.2 × 232.9	1,200	1.41
Ariel	13.4	581.1 × 577.9 × 577.7	1,670	2.52
Umbriel	11.7	584.7	1,400	4.14
Titania	35.2	788.9	1,710	8.70
Oberon	30.1	761.4	1,630	13.46
NEPTUNE'S MAJOR MOONS				
Naiad	—	29	—	0.29
Thalassa	—	40	—	0.31
Despina	—	74	—	0.33
Galatea	—	79	—	0.43
Larissa	—	104 × 89	—	0.55
Proteus	—	218 × 208 × 201	—	1.12
Triton	214.7	1,352.6	2,050	–5.88
Nereid	0.2	170	1,000	360.14

Notes: Moons are listed in order of their distance from the planet. A minus sign preceding orbital period signifies retrograde motion.

Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.

magnetic field of Jupiter and with the plasma disk (a disk of charged particles) that surrounds the planet.

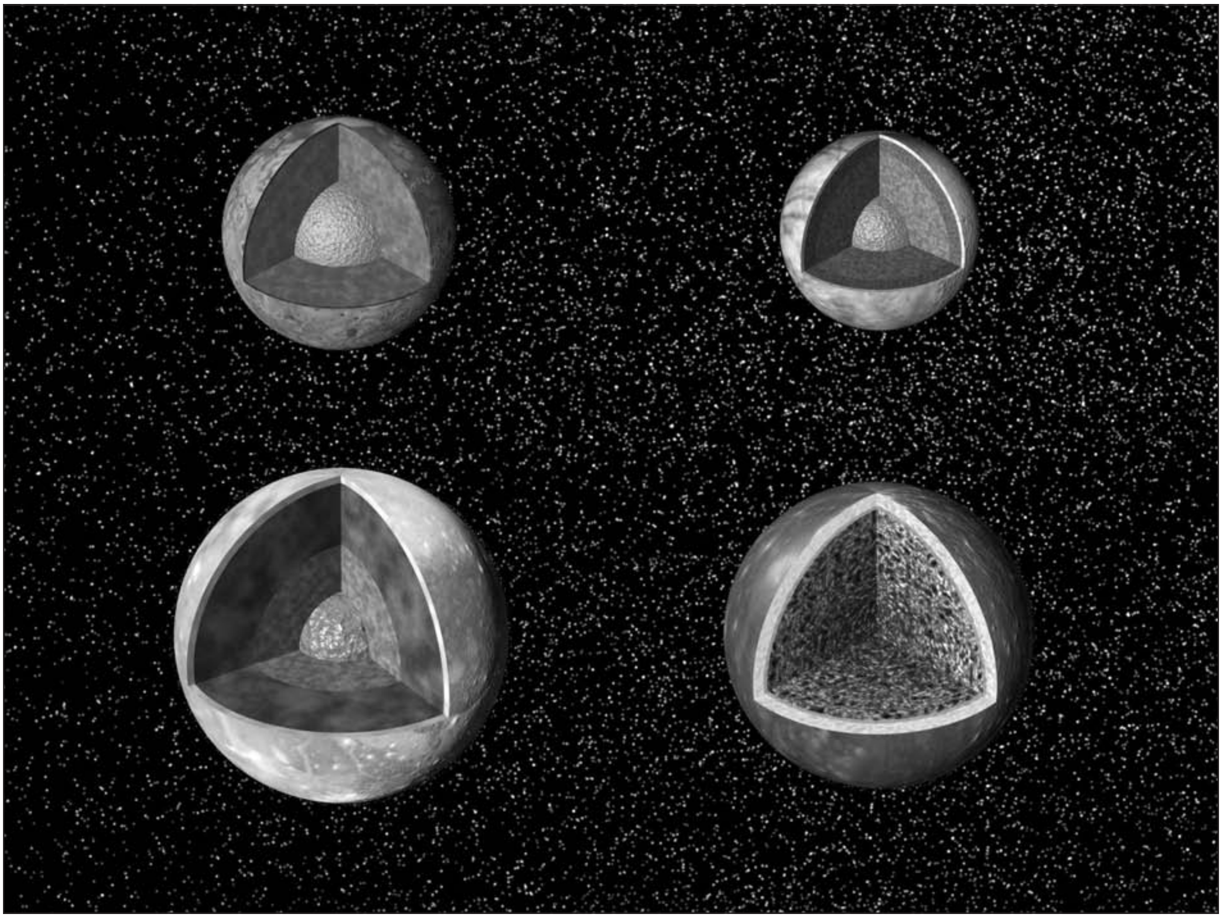
Europa is another Jovian satellite that shows signs of activity, although how recent that activity may have been is not known. Some scientists believe that there may be an ocean of liquid water as much as 100 kilometers deep lying below the icy crust of Europa. Occasional cracks in the crust allow the water to flow onto the surface, erasing any evidence of past impacts, when large flows occur, or simply to come up through the cracks to form the ridges that mark the surface.

Ganymede also shows evidence of having been active after the major cratering epochs ended within the solar system. Mountains formed from tension in the surface are evidently a consequence of a series of internal upheavals during the earlier life of the satellite. It is possible that these changes were initiated by changes in the crystal structure of the ice as the core of Ganymede slowly cooled. Resulting expansions and contractions could have cracked the surface and allowed the lower areas to be flooded with a lavalike flow of liquid water.

Callisto is believed to have been inactive since the initial formation of its crust. It is generally accepted that Ganymede and Callisto are differentiated objects. This means that initially the objects were a uniform mix of water ice and rocky materials. This mixture was heated from some source, perhaps radioactivity of the rocky material, and the denser, rocky material settled to the core, leaving the less dense water on the surface, where it eventually froze and formed the crust. For Callisto, it seems that this development brought an end to its self-generated geologic activity. All Callisto's other topographic features resulted from collisions with other space objects.

CONTEXT

Although the Galilean moons have been known since 1610, because of their small size astronomers initially found it difficult to identify details on the surface of any of them. Photographs taken by equipment carried high into



An artist's cutaway views of the interior structures of the four major Jovian satellites, as implied by gravitational and magnetic field experiments (clockwise from upper left): Io, Europa, Callisto, and the largest, Ganymede. Callisto is the only one of the four that does not show evidence of a metallic iron-nickel core surrounded by a mantle of rock. Ganymede and Europa are believed to have outer shells of water (ice or liquid). Callisto is believed to consist of a mixture of ice and rock. (NASA/JPL)

Earth's atmosphere showed Io to have dark polar regions. Careful measurements of brightnesses indicated that there was a variation as the position in orbit changed, suggesting that the surface on a given satellite did not have a uniform reflectivity. These data also led astronomers to the conclusion that the satellites were tidally locked to Jupiter, so that their periods of rotation were equal to their periods of revolution.

Another interesting feature of the orbits is that the periods of revolution of Io, Europa, and Ganymede are tied together by gravitational coupling. The period for Europa is twice that for Io, and the period for Ganymede is twice that for Europa. If any one of these three orbits changed, the other two would also change to restore the ratios that currently exist. One difficulty in the models that describe the origin of Jupiter is that they offer no account of a process through which these three satellites could move into such an orbital relationship.

Voyager data revealed four satellites that were different from expectations in many ways and that had significant differences among themselves. Models have been constructed that attempt to describe processes that would lead to the formation of the moons and give rise to their current conditions. The region in which Io and Europa are believed to have been formed would have had a fairly high temperature, and volatile materials such as water would not have been present in significant quantities. The tidal heating that Io experiences has driven off any volatile materials, such as water and carbon dioxide, that may have been a part of the original body.

Europa retained more water at formation than did Io and has not lost it. There appears to be enough heat in the silicate core to keep much of the water in the liquid phase and to allow it to flow through cracks that form in the moon's surface. This flow process has essentially resurfaced Europa during its history.

Ganymede has a mass that is roughly 50 percent water. Much of this water is in liquid form. It is assumed that the core is warm enough to transfer energy to the water to keep it liquid. The water's liquid state allows parts of the icy surface to be renewed.

Callisto's icy surface is covered with significant quantities of silicate materials. Although there may be a layer of liquid water in the interior, the crust is thick enough to prevent any of the liquid from reaching the surface. The process of plastic deformation has eliminated the features of the larger impact basins, although some of the materials that would have been part of the basin are visible in the icy crust, giving the impression that the moon has a large bull's-eye drawn on it.

The next stage in the investigation of Jupiter's satellites would be to send to Jupiter a spacecraft capable of performing detailed investigations over prolonged periods in close orbital proximity to the satellites rather than just doing multiple flybys. Such a mission was proposed and initially funded. The Jupiter Icy Moons Orbiter (JIMO) would have used nuclear propulsion not only to enter the gravitational sphere of Jupiter but also to go from orbit around one satellite to another over a very prolonged mission lasting perhaps a decade. When the program met technical difficulties and costs threatened to rise well above the allocated $1 billion, JIMO was canceled. Nevertheless, among planetary scientists, an orbiter around Europa and perhaps other Galilean moons remains a high priority for understanding the Jovian system.

Dennis R. Flentge

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. Cambridge, England: Cambridge University Press, 2007. A comprehensive work about the biggest planet in the solar system. A series of articles provided by recognized experts in their field of study. Excellent repository of photography, diagrams, and figures about the Jupiter system and the various spacecraft missions that unveiled its secrets.

Bredeson, Carmen. *NASA Planetary Spacecraft: Galileo, Magellan, Pathfinder, and Voyager*. New York: Enslow, 2000. This book is part of Enslow's Countdown to Space series. Provides an overview of NASA planetary exploration during the last two decades of the twentieth century. Suitable for all audiences.

Cole, Michael D. *Galileo Spacecraft: Mission to Jupiter*. New York: Enslow, 1999. Provides a full description of the Galileo spacecraft, its mission objectives, and science returns through the primary mission. Particularly good at describing mission objectives and goals. Suitable for a younger audience.

Davies, Ashley Gerard. *Volcanism on Io: A Companion with Earth*. Cambridge, England: Cambridge University Press, 2007. Provides the full analysis of Io's diverse set of active volcanoes based on all spacecraft data and remote sensing from Earth. Technical but accessible to the general science reader too.

Fischer, Daniel. *Mission Jupiter: The Spectacular Journey of the Galileo Spacecraft*. New York: Copernicus Books, 2001. Suitable for a wide range of audiences. Thoroughly explains all aspects of the science and engineering of the Galileo spacecraft. Particularly good are the discussions about the nature of the Galilean satellites.

Geissler, Paul E. "Volcanic Activity on Io During the Galileo Era." *Annual Review of Earth and Planetary Sciences* 31 (May, 2003): 175-211. The definitive work describing the physics and planetary geology of volcanoes on Io. Provides a complete picture of Voyager and Galileo spacecraft results.

Greenberg, Richard. *Europa the Ocean Moon: Search for an Alien Biosphere*. Berlin: Springer, 2005. A complete description of knowledge of Europa through the post-Galileo spacecraft era. Discusses the astrobiological implications of an ocean underneath Europa's icy crust. Well illustrated and readable by astronomy enthusiasts and college students.

Harland, David H. *Jupiter Odyssey: The Story of NASA's Galileo Mission*. New York: Springer Praxis, 2000. Provides virtually all of NASA's press releases and science updates during the first five years of the Galileo mission in a single work. Provides a preview of the Cassini mission. An enormous number of diagrams, tables, lists, and photographs.

Leutwyler, Kristin, and John R. Casani. *The Moons of Jupiter*. New York: W. W. Norton, 2003. Casani was the original Galileo program manager. This volume offers a heavily illustrated discussion of the Galilean satellites and a number of the less well known Jovian moons. The authors attempt an artful text accompanying the scientific findings, which may or may not be to the taste of all readers.

Lopes, Rosaly M. C., and John R. Spencer. *Io After Galileo: A New View of Jupiter's Volcanic Moon*. Heidelberg: Springer, 2007. Another member of the Springer Praxis Space Exploration series, this book summarizes the knowledge gained by Galileo about Io. Suggests new investigations needed to explain those questions that remain about the volcanism of this Jovian moon. Technical.

See also: Callisto; Enceladus; Eris and Dysnomia; Europa; Ganymede; Iapetus; Io; Jovian Planets; Jupiter's Atmosphere; Jupiter's Great Red Spot; Jupiter's Interior; Jupiter's Magnetic Field and Radiation Belts; Jupiter's Ring System; Miranda; Planetary Ring Systems; Planetary Satellites; Titan; Triton.

K

Kuiper Belt

Categories: Small Bodies; The Solar System as a Whole

The observation of object 1992 QB led to the discovery of a vast, previously unknown region of the solar system, the Kuiper Belt, which is thought to be the source of most short-period comets.

OVERVIEW

Two lines of evidence led scientists to suppose that the Kuiper Belt exists. First, comets that come close enough to the Sun to grow spectacular tails are melting, and so they cannot last forever. Furthermore, if they approach too close to the Sun (or the gas giant Jupiter), gravity may pull them apart, as it did Comet Shoemaker-Levy 9, whose life ended when it impacted Jupiter in 1994. A comet can survive only a limited number of close approaches to the Sun, so why, scientists wondered, were there still so many comets? Where did they come from?

The basic answer is the Oort Cloud, a hollow, spherical cloud of comets with the Sun at its center, extending nearly halfway to neighboring stars. It is generally accepted that long-period comets (those with periods greater than two hundred years) come from the Oort Cloud, and that, if they come close enough to Jupiter to be affected by Jupiter's gravity, some can become short-period comets (those with periods less than two hundred years).

Another feature of comets is that long-period comets come toward the Sun from all directions; that is, their orbital planes may make large angles with the ecliptic plane, the plane of the Earth's orbit around the Sun. It is called the ecliptic because when the Moon is also in this plane, an eclipse can occur. The orbits of long-period comets can be explained if they come from the Oort Cloud, and if the Oort Cloud is spherical. The orbits of short-period comets, however, are more nearly in the ecliptic plane, so their source ought therefore to lie more nearly in the ecliptic plane.

A second line of evidence for the Kuiper Belt comes from modeling the formation of the solar system. There should have been 40 or 50 Earth masses of icy-rocky material beyond Neptune. What happened to that material? In 1949, Kenneth Edgeworth speculated that a reservoir of comets existed beyond the orbit of Neptune. Then, in 1951, Gerard Kuiper reasoned that there should have been large amounts of icy material beyond Pluto and that Pluto's gravity should stir up the icy bodies that formed there, flinging some of them sunward as short-period comets and flinging others into interstellar space. It was thought then that Pluto's mass was far larger than it actually is.

In 1980 Julio Fernández published a paper that used the words "Kuiper" and "comet belt" for the first time in one sentence. Subsequent researchers combined the words, and the term "Kuiper Belt" was born. According to Fernández, a belt of potential comets lies beyond the orbit of Neptune in the ecliptic plane. This belt provides the short-period comets and repopulates the Oort Cloud. Fernández estimated that the Oort Cloud lost three hundred comets per year as they were flung inward toward the Sun or outward to interstellar space. Therefore, the Oort Cloud could not endure for the life of the solar system unless it was being repopulated.

Pluto's strange properties offer another clue that there ought to be a Kuiper Belt. In 1977 James Christy discovered the large satellite of Pluto subsequently named Charon. Using Charon's orbital period and its separation from Pluto, scientists could calculate the masses of both Pluto and Charon. Pluto's mass is 0.0021 Earth mass, only one-sixth the mass of Earth's moon. Charon's mass is 0.0003 Earth mass. Pluto's orbit does not lie in or near the ecliptic plane, as do the orbits of the major planets. Charon was most likely formed in a collision between Pluto and another object perhaps one-fifth the size of Pluto. For this to have any likelihood, there must have been many such objects present, far more than seem to be around today.

The discovery of the Kuiper Belt object (KBO)

2003 UB_{313}, later named Eris, caused astronomers to reevaluate the definition of a planet, since Eris is about 25 percent larger than Pluto. A parallel situation occurred when the first asteroids were discovered and included in the list of planets until it was decided that they belonged to a different class of objects. Likewise, Pluto is not a planet, but the first of the KBOs to be discovered. Pluto and Eris are the largest of the known KBOs and are now classed as dwarf planets, along with the largest asteroid, Ceres.

Also in 1977, Charles Kowal discovered an object that was eventually named Chiron, for the mythological centaur of that name. Chiron is estimated to be 170 kilometers (106 miles) in diameter and orbits between Saturn and Uranus. Apparently Chiron is a comet, because on various occasions it has produced a coma (the large vapor cloud typically identified as the "tail" of a comet). About a hundred such objects have been found with orbits that cross the orbit of at least one of the giant planets. All such objects are referred to as centaurs, and at least three of them have produced comas. Because of gravitational tugs from the giant planets, centaur orbits are unstable over periods of millions of years, so the centaurs now seen must have migrated to their present locations during the past 10 million years. The centaurs will eventually either crash into one of the giant planets or migrate to stabler orbits.

KNOWLEDGE GAINED

The centaurs and KBOs can be grouped by color: Some are red, and others are blue-gray. The red color can arise as cosmic rays and solar ultraviolet rays strike carbon compounds on an object's icy surface to form reddish-brown compounds. The blue-gray color may arise as smaller objects strike and crater a KBO, covering its surface with new layers of ices. Based on analyses of comets, KBOs must be largely water ice along with carbon dioxide ice, carbon monoxide ice, methane ice, methanol ice, ammonia ice, amorphous (noncrystalline) carbon, silicates and other stony materials, sodium, carbonates, simple hydrocarbons, and clays. The Deep Impact mission found that Comet 9P/Tempel was covered by a dust layer tens of meters deep and the nucleus of the comet was 75 percent empty space, which made it structurally weak. The minor constituents in particular may differ from one comet to another, implying that comets formed at various distances from the Sun.

Experimental evidence of the Kuiper Belt came with the 1992 discovery by Jane Luu and David Jewitt of object 1992 QB_1. It was 41 AU from the Sun, which is about the orbit of Pluto. They estimated its diameter at 250 kilometers. Six months later they found 1993 FW, a second candidate for the Kuiper Belt. By early 2008, more than one thousand KBOs had been discovered. The main belt is shaped like a doughnut centered on the Sun. The doughnut itself goes from 30 AU from the Sun to perhaps 100 AU or even 1,000 AU or more, although there is a sharp falloff in numbers at 50 AU. Objects in the belt are called CKBOs, for "classical Kuiper Belt objects," also called cubewanos, named for the pronunciation of the first CKBO discovered, 1992 QB_1 (Q-B-1-0-s). In order not to be influenced by Neptune at 30 AU from the Sun, these objects must have average distances from the Sun of 40 and 50 AU.

A large fraction of known KBOs are plutinos. Like Pluto, plutinos make two trips around the Sun for every three trips made by Neptune. This guarantees that if they are not now close to Neptune, they will not be close in the future, and therefore their orbits are stable. Since they never get close to Neptune, they may approach to within 30 AU of the Sun with impunity. Pluto never comes closer to Neptune than 17 AU, even though Pluto's orbit crosses Neptune's orbit.

The final group of Kuiper Belt objects are scattered disk objects, or SDOs. Their orbits tend to be more elliptical than the classical Kuiper Belt objects' orbits, and they also travel considerably above and below the ecliptic plane. Scattered disk objects and centaurs seem to form a continuous distribution. If a KBO is ejected to the inner solar system, it becomes a comet or a centaur. It is estimated that there may be 70,000 KBOs between 30 and 50 AU with diameters greater than 100 kilometers (60 miles), and perhaps 1 million with diameters of 1 kilometer (0.6 mile) or larger. There may be as many as 30,000 SDOs that are 100 kilometers (62 miles) in diameter.

CONTEXT

It is believed that stars like our Sun form from a more or less spherical cloud of gas and dust within a larger cloud. When conditions are right for one star to form, they are often right for many stars to form; hence, stars, like puppies, tend to be born in litters. The close approach of one or more neighboring stars may have sheared off the Sun's Kuiper Belt at about 50 AU from the Sun. Solid material left over from the formation of the Sun would have developed into a rotating disk with the newborn Sun at its center. Concentrations in the disk formed into asteroid-sized bodies that in turn combined to form the planets. Icy asteroids would have formed among the giant planets and beyond. Those beyond Neptune became the Kuiper Belt. In the inner belt, from about 30 AU to perhaps 50 AU, collisions occasionally occurred, with the result that large objects grew larger (such as Pluto and Eris) and small objects were eventually ground to dust. Beyond 50 AU, objects (if they exist) should be pristine samples of the solar nebula, since they would be so far apart that collisions would be unlikely.

As the newly formed planets interacted with icy asteroids around them, some icy bodies would have been flung inward toward the Sun, and some of these may have hit the Earth and formed the oceans; the Earth was probably born with far less water than it now has. Others were flung outward to form the Oort Cloud and perhaps scattered disk objects. Interactions between KBOs today should occasionally propel them out into the Oort Cloud or inward as comets in the inner solar system or as centaurs in the outer solar system.

Some questions about Pluto and other KBOs may be answered by the New Horizons mission, sponsored by the National Aeronautics and Space Administration (NASA). Launched in 1995, it should reach Pluto by 2015. After taking images in the visible, infrared, and ultraviolet bands of Pluto and its satellite Charon, the spacecraft will go on to one or more KBOs. The infrared images should reveal the surface composition of these objects.

Charles W. Rogers

FURTHER READING

Davies, John. *Beyond Pluto: Exploring the Outer Limits of the Solar System*. New York: Cambridge University Press, 2001. Excellent and easy-to-read discussion of the ideas and observational evidence leading up to the discovery of the Kuiper Belt. Describes the properties of the various classes of objects found.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory text covering the field of astronomy. Contains descriptions of astrophysical questions and their relationships.

Lin, Douglas N. C. "The Chaotic Genesis of Planets." *Scientific American* 278, no. 5 (May, 2008): 50-59. Discusses the formation of the solar system, including the formation of the Kuiper Belt and the Oort Cloud.

Luu, Jane X., and David C. Jewitt. "The Kuiper Belt." *Scientific American* 274, no. 5 (May, 1996): 46-52. This landmark article is accessible to the layperson. Luu and Jewitt discuss how they found KBOs and helped to establish the existence of the Kuiper Belt.

Malhotra, Renu. "Migrating Planets." *Scientific American* 281, no. 3 (September, 1999): 56-63. During the formation of the solar system, Jupiter migrated inward, while Saturn, Uranus, and Neptune migrated outward by slinging icy bodies to the inner solar system and to the Kuiper Belt and Oort Cloud.

Stern, S. Alan. "Into the Outer Limits." *Astronomy* 28, no. 9 (September, 2000): 52-55. Discusses history of the Kuiper Belt and how Pluto fits as a KBO.

See also: Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Interplanetary Environment; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Nemesis and Planet X; Oort Cloud; Pluto and Charon; Solar System: Origins.

L

Life's Origins

Category: Life in the Solar System

At some time after the formation of the Earth some 4.6 billion years ago and before the earliest known microfossils, some 3.5 billion years old, life evolved from widely available organic compounds into unicellular creatures similar to modern blue-green algae. Plausible schemes for their evolution from abiotically produced organic matter remain elusive. Much current research into the origin of life is based on the idea that the environment of early life was as important as its nature and compares the relative merits of the ocean, submarine volcanic vents, and soil as possible places where life began.

OVERVIEW

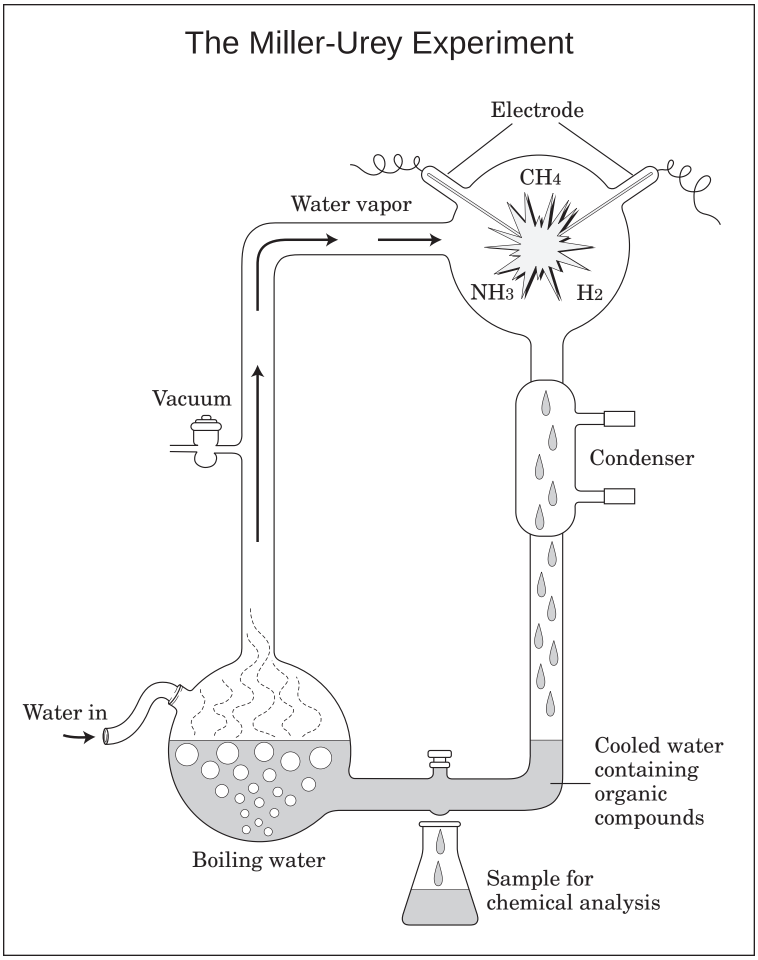
Life on Earth is no older than the solar system, which formed some 4.5 billion years ago. Early stages in the accretionary process by which Earth formed are revealed by meteorites, most of which have been radiometrically dated to about this age. A particular kind of meteorite called a carbonaceous chondrite is characterized by the high-temperature globular mineral grains called chondrules, mixed with low-temperature iron-rich clays and organic compounds. Most of the organic compounds are dispersed among the clays, but some of them form organic lumps and globules.

These latter have been mistakenly identified as microfossils, but the organic compounds of carbonaceous chondrites have some diagnostic features of organic matter formed abiotically. Many organic molecules, such as sugars, are asymmetric. When synthesized in the laboratory, there are about equal amounts of left-handed and right-handed versions (each is known as an enantiomer) that polarize light in a clockwise or counterclockwise direction, respectively. Organic matter in carbonaceous chondrites contains similar proportions of left-handed and right-handed molecules. However, living organisms are highly selective for just one enantiomer of a compound, including, for example, only D enantiomers of sugars. Organic compounds that also are presumed to have formed abiotically have been detected from the characteristics of light spectra reflected from asteroids, comets, and interstellar dust. Thus, simple organic compounds used by living creatures are widely available in the universe and have been for at least 4.5 billion years.

It is an enormous step from a tarry mixture of simple organic compounds to a fully functioning organism, with its long polynuceotide instructions and messengers and its complex molecular assembly line and controlling compounds of protein all bound up within a system of membranes and cell walls. This degree of complexity, seen in some of the simplest living forms of unicellular life, was attained by about 3.5 billion years ago. Microscopic fossils of unicellular cyanobacteria (blue-green algae) have been found permineralized in black cherts of this age in the Pilbara region of Western Australia. These are among the oldest little-deformed sedimentary rocks in which such fossils are likely to be found. However, carbon compounds found in some highly deformed and metamorphosed sedimentary rocks at least as old as 3.8 billion years (and possibly some more than 4 billion years old) yield carbon isotope ratios like those found in photosynthetic life, suggesting life may have existed back then.

Older fossil evidence of life on Earth is unlikely. There is, however, a younger fossil record of increasingly complex microfossils in cherts and shales ranging from 3.5 billion years ago to the present. By 700 million years ago, there were also complex multicellular creatures. Fossil plants and animals become increasingly familiar in rocks of younger geological age.

The vast gulf between primordial organic compounds and functioning cells revealed in the rock and meteorite record also is apparent from



It has been shown many times that organic compounds, the beginnings of life, including amino acids, are produced readily within water in sealed flasks that contain reducing gases such as carbon dioxide, if these are energized by electrical discharges, ultraviolet light, or even shock waves. Shown here is the famous Miller-Urey experiment of 1955, in which such organic compounds were produced. With the right conditions, it is conceivable that similar reactions may have occurred on other planets or their satellites—not only in our own solar system but in other star systems as well.

experimental studies in organic chemistry designed to investigate how life originated. It has been shown many times that organic compounds including amino acids are produced readily in sealed flasks containing water and reducing gases such as carbon dioxide energized by electrical discharges, ultraviolet light, or even shock waves. The exact gas mixture or energy source is not critical to organic matter production, although yields are limited in gas mixtures including oxygen or with weak energy sources. Clay, siderite, and other minerals added to these reaction flasks have been shown to promote both the production of simple organic compounds and their assembly into larger compounds. How they accomplish it is not certain: They may act as a template to hold molecules in a favored orientation for combination, either as a catalyst participating in reversible reactions that tend to promote a particular combination or as a chemostat in buffering solutions from extreme acidity. Thus, the origin of organic compounds such as those in carbonaceous chondrites can to some extent be duplicated in a chemistry laboratory.

Chemical experiments fall far short of creating such complex organic molecules as ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). There have been cases where some parts of nucleotides have been synthesized and where nucleotides have been aggregated into polynucleotides. It remains a problem how there could arise a nucleotide-rich environment of the correct chemical nature under abiotic conditions. Nevertheless, it is clear from much experimental work in molecular biology that polynucleotides such as RNA and DNA are the basis for a system of information storage and transfer in all known lifeforms on Earth.

The odds of tarry abiotic organic matter becoming organized into a functioning cell by pure chance are astronomically unfavorable. It has been argued that, given the long period of time available, even the rarest of chances becomes a certainty. This argument now has little appeal, because geological and meteoritic evidence constrains the origin of life to within the first billion years of Earth's history and because studies in molecular biology have revealed astounding complexity in even the simplest of organisms. One escape from this dilemma is to postulate seeding of the Earth with some form of extraterrestrial life, either by long-distance transport of spores across space or by the technological enterprise of an extraterrestrial civilization. These views have not proven very productive, because they are difficult to test and, in addition, do not solve the problem. Even the simplest life-forms are well suited to Earth and are made of materials commonly available here. Thus, if life is extraterrestrial, it originated under broadly Earthlike conditions elsewhere in the universe.

It remains a useful exercise to speculate on how life might have arisen from natural causes here on Earth. The most widely held hypothesis for the origin of life until recently has been that life arose in the primordial soup of the world ocean. The early atmosphere of the Earth probably included little or no free oxygen, like that of most planets. Under these conditions, organic matter would have been produced in the sea by lightning strikes, ultraviolet light from the Sun, and shock waves from earthquakes or meteorite impacts. This vast organic nutrient solution is viewed as the source of ever more complex molecules, some of which attained the critical innovation of self-replication. These early versions of RNA and DNA then evolved increasingly complex mechanisms for self-preservation, such as cell organelles, cells, and bodies. By this view, the primitive molecular ecosystem fed on abiotically produced organic compounds, and the earliest organisms may have been fermenters. It was only when primordial organic compounds of the ocean were depleted that there arose more sophisticated cellular machinery for the biological production of organic matter by photosynthesis.

The most serious problem of this hypothesis is the gap between abiotic organic matter and organized self-replicating molecules. The ocean is, and probably was, too dilute a solution, too low in dissolved phosphorus, and too uniform in its alkalinity and oxidation state for synthesis and maintenance of complex organic molecules. The laboratory synthesis of DNA requires many different reactions under very different chemical conditions: acidic and alkaline, oxidizing and reducing, wet and dry. Such fluctuating conditions can be imagined during the intermittent evaporation of seawater ponded on ice floes or in tidal flats, but low temperatures and the formation of salts would retard production of complex organic molecules.

The discovery of deep-sea volcanic vents prompted a second group of hypotheses about the origin of life. These "black smokers" are openings along mid-ocean ridges where internal water of the Earth as well as seawater percolating through seafloor basalt has been heated by contact with hot volcanic rocks and is vigorously expelled into the sea at temperatures of up to about 380° Celsius (about 650 kelvins). The water gushing from the vents is also strongly acidic and depleted in oxygen. The "black smoke" billowing forth, as well as the vent chimneys, consists of dark sulfide minerals. The idea that ocean vents are flow reactors capable of stripping oxygen from carbon dioxide to produce organic matter is one of the arguments that they may prove pertinent to the origin of life. In addition, the complex system of fractures and vesicles in volcanic rocks around vents would have encouraged a variety of simultaneous and very different chemical reactions, in contrast to the tendency for chemical uniformity in the open ocean. Furthermore, experimental studies have shown that proteins aggregate into spheres and sheets at high temperatures, and such structures can be imagined as crude membranes and cell walls. If life did evolve in oceanic vents, then it would initially have lived by fermentation or respiration of abiotically produced organic compounds, and photosynthesis would have arisen with more complex organic systems at those few vents exposed to sunlight on land or in shallow water.

Despite some appealing features of this hypothesis, it remains to be demonstrated whether the minute amounts of organic matter detected in submarine volcanic vents were produced abiotically or are remains of the distinctive fauna of tube worms, clams, and crustaceans that now colonize the seafloor around vents. The high temperatures and strongly acidic nature of vent water are incompatible with the production and preservation of complex proteins and nucleotides. These problems are reduced considerably even a short distance away from the vent, but then there remain the problems already mentioned for an open ocean origin of life.

A third set of hypotheses, that life evolved in soil, can be taken to include ideas involving clay minerals. Soils are, and probably were, the principal sites where clay is formed on Earth. Soil hypotheses for the origin of life combine many of the desirable features of ocean and vent hypotheses. Soil water could have been enriched in organic matter like the early ocean, and soils are in a sense flow reactors driven by the energy of the Sun at the surface. Soil water is compartmentalized and evaporates into a complex array of films around mineral grains and other particles. Soils enjoy moderate temperature and mild chemical conditions.

The most critical advantage of the soil hypothesis is that the production of clay by weathering or of organic matter by abiotic synthesis would have been encouraged because clayey and organic soils are less easily eroded than are sandy, inorganic soils. This matter of advantage can be imagined as a crude form of "natural selection" in which the most clayey and organic soils survived erosion for the longest periods of time. Thus, even the least efficient mechanisms of abiotic weathering and organic synthesis could result in soils as organic and clayey as carbonaceous chondrites. Increasingly complex compounds such as sugars and proteins would be preserved preferentially to the extent that they could bind the soil and maintain it within the zone of energy transfer at the surface. In this view, the primary production of organic matter would initially have been fueled by light from the Sun at the surface, and photosynthetic organisms may well have preceded fermenting and respirating creatures. The idea that life arose in soils reduces some of the difficulties of forming functioning cells from tarry organic matter, but it cannot be claimed to eliminate them.

METHODS OF STUDY

Clues to the origin of life might be found from the study of Precambrian microfossils. Their fossil record is known back to the base of the useful sedimentary record, and the nature of the most ancient microfossils provides constraints on the origin of life. Most of the microfossils have been found in silicified limestones of a distinctive domed and banded kind called stromatolites and in black, organic-rich cherts. The rocks are cut and polished as thin, petrographic sections for observation under the microscope. Such studies also can be supported by isotopic and chemical research on the organic matter preserved in the chert.

Other constraints on the origin of life come from the study of meteorites, particularly carbonaceous chondrites. They are relatively rare because they are rapidly destroyed by weathering after they land on Earth. However, increasing numbers of them have been found in Antarctica, where little weathering occurs and the dark meteorites stand out starkly against the white snow and ice. They can be studied in thin sections and by an array of methods for mineralogical and chemical characterization of their constituents: X-ray diffraction of their clays, transmission electron microscopy of their minerals and organic structures, electron microprobe chemical analysis of their minerals, and mass spectrometric separation and identification of their organic compounds.

Much useful information on the origin of life also has come from laboratory experiments on the synthesis of organic matter, on the biochemistry of living organisms, and on the nature of clays and other minerals. Organic synthesis experiments involve placing water and mixtures of gases in a sealed flask and applying an energy source, such as an electrical spark. The resulting tarry organic matter can be analyzed by a variety of methods, including spectrometric separation and identification of organic compounds.

A major enterprise of molecular biology is working out the structure of the various complex organic molecules found in living creatures, using such methods as X-ray diffraction, nuclear magnetic resonance, and mass spectrometry. Molecular biologists also strive to understand how the molecules may work from experiments with simple organisms such as bacteria and viruses.

CONTEXT

Few topics in science are as inaccessible to direct observation or as fraught with philosophical and religious implications as the origin of life. Scientific techniques are not designed to answer such questions as why humankind exists, but evidence from astronomical, geological, chemical, and biological sciences is revealing some possibilities of how life may have originated. It is doubtful that any of the hypotheses concerning the origin of life will ever become so firmly established as are the laws of physics. The long-accepted concept of molecular evolution in the primordial soup of the world oceans is challenged by alternative views of life's origin in submarine vents and in soils. Much current work is based on the idea that the way in which life interacts with its environment is as important as its physical nature and the way in which it reproduces. Scientists may never satisfactorily reveal how life came about, but they are discovering many fundamental interrelationships between Earth and its life.

Gregory J. Retallack

FURTHER READING

Bernal, J. D. *The Origin of Life*. New York: World Books, 1967. This comprehensive introduction to the topic, though somewhat dated, is written at a level accessible to the layperson. It emphasizes the origin of life in the sea. An appendix includes a translation of the original Russian article on this idea, first published by I. A. Oparin in 1924.

Cairns-Smith, A. G. *Genetic Takeover*. Cambridge, England: Cambridge University Press, 1982. This masterpiece of logic develops the startling theory that clay not only encouraged the origin of life but also can be considered as part of a quasi-living system itself. This book assumes a university-level background in science, but care is taken to explain technical lines of evidence in various scientific specialties. Includes a comprehensive bibliography.

Crick, Francis. *Life Itself: Its Origin and Nature*. New York: Simon & Schuster, 1981. In this slim volume, a Nobel laureate in biochemistry explains, in terms easily understood by the layperson, why life is too complex a phenomenon to have developed on Earth within the limited geological time available. In arguing that the Earth was deliberately seeded with microbial life by alien civilizations, he also reviews a variety of other hypotheses for the extraterrestrial origin of life.

Hazen, Robert. *Genesis: The Scientific Quest for Life's Origins*. Washington, D.C.: Joseph Henry Press, 2005. Hazen, a researcher in the Carnegie Institution's Geophysical Laboratory and Clarence Robinson Professor of Earth History at George Mason University, presents the field of abiogenesis—the scientific discipline devoted to the investigation of life's origins, discussing not only the science but also the discoveries, dramas, and debates.

Levinton, Jeffrey S. *Genetics, Paleontology, and Macroevolution*. 2d ed. New York: Cambridge University Press, 2001. A comprehensive work on macroevolution focusing on the evolution of animals. Covers genetics, evolution, fossil lineage, and the Cambrian explosion.

Luisi, Pier Luigi. *The Emergence of Life: From Chemical Origins to Synthetic Biology*. New York: Cambridge University Press, 2006. A review of the scientific field that focuses on the chemistry of the origin of life, as well as its literature. Designed as a text for organic chemistry and other college-level Earth and life science courses.

Miller, S. L., and L. E. Orgel. *The Origins of Life on Earth*. Englewood Cliffs, N.J.: Prentice-Hall, 1974. In this well-referenced and carefully worded account aimed at university students and professional scientists, emphasis has been placed on experimental studies on the abiotic synthesis of organic compounds.

Monod, Jacques. *Chance and Necessity*. Translated by Austryn Wainhouse. New York: Alfred A. Knopf, 1971. In an extended essay aimed at a general audience, this distinguished French molecular biologist develops a critical theme for evaluating theories about the origin of life: Given the complexity of living creatures, it is hardly likely that they originated by chance; some kind of "natural selection" is needed to convert these long odds to a necessity.

Nisbet, Evan G. *The Young Earth: An Introduction to Archaean Geology*. Winchester, Mass.: Unwin Hyman, 1987. This colorful account of the basic data and controversies involving the oldest rocks on Earth is aimed at an audience with at least high school training in science. An excellent guide for the primary geological literature on these rocks and especially for the idea that life originated in and around submarine volcanic vents.

Schindewolf, Otto H. *Basic Questions in Paleontology: Geologic Time, Organic Evolution, and Biological Systematics*. Translated by Judith Schaefer. Chicago: University of Chicago Press, 1994. First published in German in 1950, this text was highly controversial at the time because of its anti-Darwinism approach. The author argues the importance of major catastrophic events and mass extinctions. In recent decades his ideas have become more widely accepted, and the basis for current research.

Schopf, J. William, ed. *Life's Origin: The Beginnings of Biological Evolution*. Berkeley: University of California Press, 2002. Eminent winners of the Oparin/Urey Gold Medal of the International Society for the Study of the Origin of Life contribute papers on chemistry, paleobiology, and astrobiology, in an effort to illuminate the origins of life on Earth. Designed for a general audience; includes a glossary of scientific terms.

Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008. This college-level textbook for introductory geology courses is well written and illustrated. The chapter on Earth's evolution through geologic time includes the origin and evolution of life.

See also: Earth's Atmosphere; Earth's Composition; Earth's Oceans; Earth-Sun Relations; Europa; Extrasolar Planets; Extraterrestrial Life in the Solar System; Habitable Zones; Interstellar Clouds and the Interstellar Medium; Jovian Planets; Main Sequence Stars; Mars: Illusions and Conspiracies; Mars: Possible Life; Mars's Atmosphere; Meteorites: Carbonaceous Chondrites; Meteoroids from the Moon and Mars; Planetology: Comparative; Search for Extraterrestrial Intelligence; Supernovae.

Lunar Craters

Category: Natural Planetary Satellites

Most lunar craters are the erosion scars of debris left over from the origin of the solar system colliding at high velocities with the surface of the Moon. Studies of sizes and time distributions of lunar impact craters allow scientists to make estimates of the same process acting on the Earth, where much of the evidence has been removed by erosion. Volcanic craters enable researchers to determine the eruption characteristics and thermal evolution of the Moon.

OVERVIEW

All of the large lunar craters are named; many of these names are attributable to a 1651 publication by Giambattista Riccioli in which they appear on a map drawn by P. Grimaldi. Riccioli divided the nearside of the Moon into octants. This map was drawn with the aid of the "Galilean" rather than the "astronomical" telescope and so was not inverted. On this map, Octant 1 extended from the ten o'clock position to just past eleven o'clock and was succeeded clockwise by the seven other octants. Craters were named for astronomers, beginning with the most ancient in Octant 1 and concluding with Riccioli's contemporaries in Octant 8. This practice has continued to the present, with the restriction that a crater is always named for a scientist no longer living.

Three principal processes have created the lunar craters. There are those directly excavated by the impact of a meteorite; there are those, called secondaries, that result from the impact of material excavated to form the crater of the primary meteorite impact; and there are those of volcanic origin. Until the return of lunar samples from the Apollo missions (1969-1972), the scientific community had been sharply divided into those who believed the majority of lunar craters to be of impact origin and those who believed the majority to be volcanic in origin. Evidence gained as a result of the Apollo missions has established that the vast majority of lunar craters are of impact origin, resulting from collisions of the Moon with meteors, asteroids, comets, and minor planets (large objects that failed to achieve independent planetary status) at velocities of from 5 to 50 kilometers per second (10,000 to 100,000 miles per hour).

No one has counted the total number of lunar craters, as they range in size from the microscopic to the giant (2,500-kilometer-diameter) South Pole-Aitken basin. It has been estimated, however, that there are about 1,850,000 craters with diameters in excess of 1 kilometer on the lunar surface and 125 with diameters greater than 100 kilometers. A 3,200-kilometer-diameter Procellarum basin has been tentatively identified which, if placed over a map of the United States, would stretch from Washington, D.C., to western Utah and from Brownsville, south Texas, up into central Canada. At the other end of the scale, microscopic impact craters are produced on the Moon because of the lack of a lunar atmosphere. Similar-sized particles rapidly burn up in the Earth's atmosphere.

Primary impact craters increase in morphological complexity with increasing size. Small craters are bowl-shaped, with a well-defined, generally circular rim, smooth interior walls, and a depth-to-diameter ratio of 1:5 to 1:6. The floor of a fresh crater is invariably at a lower elevation than the preexisting terrain. The rim of the crater is surrounded by a generally circular continuous ejecta blanket, followed outward by the discontinuous ejecta blanket. This discontinuous ejecta often takes the form of rays that radiate outward from a zone close to the center of the primary impact site. An exception is found in craters produced by highly oblique impacts. These craters are generally elongated and have ejecta blankets preferentially distributed downrange or exhibiting a bilateral symmetry, with "wings" on either side of the crater.

An abrupt change in the crater's shape takes place at a diameter of about 16 kilometers in the maria (the Moon's dark lava expanses) and 21 kilometers in the highlands. At larger diameters, craters develop terraces on the interior walls, have a generally broad, level floor interrupted by small hills and mounds, develop a central peak, have a less uniform rim elevation, and have a depth-to-diameter ratio reduced to about 1:40. Flows and ponds, which are often seen both within and exterior to these craters, are impact melts resulting from liquefaction of the impactor and target rocks. At diameters in excess of 140 kilometers, the central peak becomes modified into a centralized peak ring. At

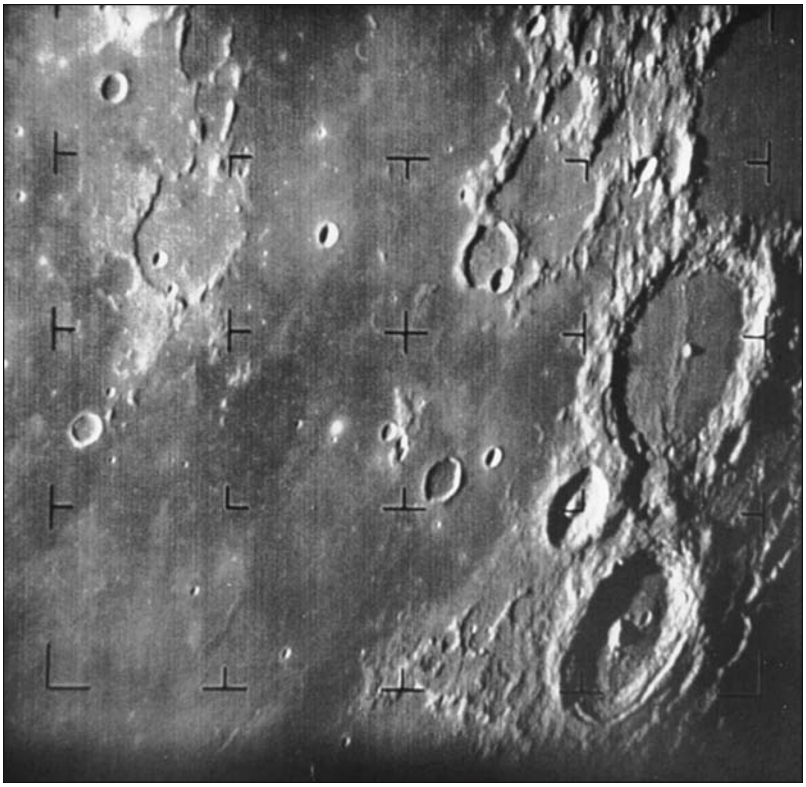
The Galileo spacecraft caught this image of the Moon in 1992, on its way to meet Jupiter in 1995-1997. Major features include Mare Imbrium (center left, Mare Seranitatis and Mare Tranquillitatis (center), Mare Crisium (right edge), and the bright Tycho basin (bottom). (NASA) diameters in excess of 350 kilometers, multiple rings of alternating elevated and depressed terrains, the giant multiringed basins, are witnessed.

Secondary impact craters are generally less regular than primaries because they are formed at lower collisional velocities, an upper limit being the lunar escape velocity of 2.4 kilometers per second, at which speed objects ejected from the surface would leave the Moon's gravitational field. The size of the secondaries is largely dependent on the size of the primary. Large secondaries have diameters between 2 and 5 percent of that of the primary. Generally, secondaries have smooth interior profiles and are shallower than primaries of the same diameter. They also differ from primaries in that their distribution is nonrandom. They frequently occur in linear or curving chains, patches, or clusters surrounding the primary. Another common feature of secondaries is the presence of a herringbone pattern produced by small ridges ploughed up by impacting objects closely spaced in both time and distance. The apex of the V shape points back toward the primary.

Many large lunar craters were once considered to be analogous to terrestrial calderas. Calderas form as a result of collapse following evacuation of a large, near-surface magma chamber. Analyses of returned lunar volcanic materials established, however, that they were derived from great depths (150-400 kilometers), with little evidence of residence at shallow levels for any extended periods. True lunar volcanic craters are recognized primarily on the basis of their distribution, which, like that of secondaries, is nonrandom. Volcanic craters or endogenic craters (those of internal origin) are found at the summits of volcanic domes and cones, at the heads of sinuous rilles, or in association with linear fractures. The craters are generally small (less than 20 kilometers in diameter) and have outlines that range from circular to elliptical to highly irregular. Some volcanic craters are surrounded by a halo of dark surface deposits believed to consist of pyroclastic materials ejected during strombolian- or vulcanian-style eruptions.

Impact craters are the product of an instantaneous geologic event, yet the lunar surface has been subjected to the formation of these features for at least the last 4.2 billion years. Many lunar craters have thus become highly modified from their original pristine form. Since the Moon lacks an atmosphere, this modification results primarily from two agents: later impacts and volcanism. An impact has two principal effects. First, it will result in the total or partial obliteration of any crater smaller than itself that was located within the area of the younger crater. Second, ejecta from the younger crater will erode the walls and infill the floors of craters surrounding itself. At the extreme, the ejecta deposits could totally infill the preexisting surrounding craters. The net result of this process is that older craters are shallower than are newer ones of similar size. The effects of volcanism on impact craters are largely restricted to areas around and within the major maria. A commonly held misconception is that an impact event can trigger the release of magma from the lunar interior. Although the relationship between depth and diameter of the large basins is a subject of much debate, there are very few who believe that the original impact cavities extended to a depth greater than that of the lunar crustal thickness of 75 kilometers. Moreover, there is good evidence that volcanism within any single large basin extended over a time frame of several hundred million years—which is difficult to reconcile with an instantaneous impact event. The reason for the association of volcanic material with impact craters is that the crater floors are topographic lows and closer to the part of the interior from which the volcanic materials were derived.

One feature attributed to volcanic modification of impact craters is the presence of floor fractures. These features are primarily found on the level floors of craters with diameters of 30 to about 100 kilometers and consist of radial and concentric arrangements of fractures resembling spiderwebs. They are attributed to uplift of the crater floor by subfloor intrusions of magma. Some of these magma bodies found outlets to the surface and, with limited volcanic output, resulted in the formation of volcanic dark-halo craters aligned along the floor fractures. The crater Alphonsus is a typical example. With more extensive volcanism, the floor of



The first close-up of Earth's moon, taken on July 31, 1964, from Ranger 7. The three large craters on the right are (top to bottom) Ptolemaeus (obscured in shadow), Alphonsus, and Arzachel. Mare Imbrium, pockmarked with smaller craters, fills most of the screen from center to left. (NASA/JPL)

the crater becomes flooded with lava until even the central peak becomes buried. At this point, there is too thick an overlying lava pile, and the magma seeks alternative routes to the surface around the periphery of the crater. Some craters have experienced postflood floor fracturing, which results from the sinking of the dense, thick lava pile or posteruptive intrusion, leading to renewed uplift. Flooding of the larger basins appears to have begun within the central low and later extended to the topographic lows between the mountain rings. Small impact craters were constantly being created during the period of basin filling. Many of these craters were either partially or totally covered by younger lava flows.

METHODS OF STUDY

Lunar crater studies began in earnest with the availability of the first crude telescopes. At this point it was realized that the Moon's surface was not perfectly smooth, as the Greeks had hypothesized, but instead was pockmarked with features that ranged in size from depressions barely resolved in the telescope to enormous basins. During this telescopic era of investigation, most scientists believed that lunar craters were formed volcanically. However, Eugene Shoemaker of the U.S. Geological Survey became a champion of an impact origin for the majority of the Moon's craters about the time that the space age dawned and early probes began to be sent to investigate the Moon at close range.

Early Pioneer probes (from the late 1950's to the early 1960's) largely failed to achieve lunar goals, but Soviet Luna and American Ranger probes purposely crash-landed on the Moon in several sites and transmitted images up until nearly the instant of impact. Those pictures revealed that the lunar surface displayed cratering down to the smallest scale, providing further evidence for their impact origin, with an enormous number of secondary craters formed out of each primary impact. Several U.S. Surveyor spacecraft (in the mid-1960's) and Soviet Luna spacecraft (from the mid-1960's to 1976) soft-landed on the lunar surface to provide far more images of the nature of the heavily cratered surface, even in the mare regions that appear relatively smooth from backyard telescope views. The Surveyors were followed by Lunar Orbiter and more Soviet Luna spacecraft, which provided detailed maps of cratering across the lunar surface, thereby assisting scientists in both the United States and the Soviet Union to identify safe landing sites to which they could dispatch astronauts and cosmonauts.

Although the Soviet Moon landing program failed, between July, 1969, and December, 1972, six Apollo missions landed at six different sites, returning samples that bore evidence of both impact origin (the breccias) and volcanism (the igneous rocks). In 1976, Luna 24 returned samples to Russia robotically. Then interest in directly investigating the Moon waned for nearly two decades.

A nearly two-decade drought in lunar exploration ended with the Clementine spacecraft, launched on January 25, 1994, by the National Aeronautics and Space Administration (NASA) and the Ballistic Missile Defense Organization. Although principally a test of new sensor technologies and spacecraft systems, Clementine used the Moon as its target and provided new insights into the Moon's surface distribution of chemicals and minerals. Clementine examined the lunar surface in several different bands of the electromagnetic spectrum, created a laser altimeter-generated map of lunar stratigraphy, and provided a more detailed gravity map of the Moon before its mission ended in June, 1994. Clementine's most exciting finding was that there were likely to be deposits of water ice on the Moon inside permanently shadowed craters, enough water to support human outposts with both potable water and fuel (by breaking down water into hydrogen [fuel] and oxygen [oxidizer]). Clementine's water detection was indirect, in that its sensors picked up the presence of hydrogen (protons), which was then interpreted to be bound in water ice.

Clementine was followed by Lunar Prospector, which launched on January 7, 1998, and conducted a 570-day examination of the lunar surface with alpha particle, neutron, and gamma-ray spectrometers. Lunar Prospector verified the signature seen by Clementine and provided a better estimate of the amount of water ice that might be available in shadowed craters on the Moon. Lunar Prospector was directed to impact the lunar surface on July 31, 1999, in an area (the shadowed crater Shoemaker) where water ice was expected to be found. It was hoped that the impact process might liberate the water ice so that it could be detected from Earth. The final experiment was disappointing: No water plume was seen.

International interest in the Moon increased in the early twenty-first century. The European Space Agency sent the Small Missions for Advanced Research in Technology 1 (SMART 1) spacecraft to the Moon using an advanced ion engine that took thirteen months to reach lunar orbit. From orbit, SMART 1 examined the surface with X-ray and infrared sensors to search for frozen water near the Moon's south pole. Until the mission ended on September 3, 2006, SMART 1 also provided high-resolution optical images of the entire lunar surface, not just craters where scientific interest in ice deposits was centered. SMART 1's mission ended with a purposeful impact at 2 kilometers per second in another attempt to kick up a water plume; results were not definite, although an impact flash was observed from Earth.

Japan launched Selene on September 14, 2007. Two weeks later it entered a highly elliptical initial polar orbit, which was later adjusted to a circular orbit just 100 kilometers above the surface. Its mission was to help set the stage for returning humans to the Moon. At this point NASA, the Chinese, Russians, Japanese, and Indians had expressed, with varying degrees of commitment, plans to land humans on the Moon to establish a permanent outpost. NASA's plans originated with the Bush White House in the aftermath of the *Columbia* accident. George W. Bush announced in early 2004 a Space Vision for Exploration policy which directed NASA to complete the International Space Station by 2010 and then retire the shuttle fleet in favor of returning astronauts to exploration beyond low-Earth orbit. The plan calls for returning astronauts to the Moon to stay by about 2020, and eventually sending humans to Mars and beyond.

The Chinese stated that the goal of its fledgling crewed Shenzhou space program was to send their taikonauts (the Chinese term for astronauts) to the Moon. As a first step the Chang'e 1 spacecraft was launched on October 24, 2007, to survey lunar craters near the South Pole that might be suitable for humans to begin constructing a base of operations for exploration of the Moon as a whole. India launched its Chandrayaan probe on October 22, 2008, entering lunar orbit on November 8. NASA was preparing to launch the Lunar Reconnaissance Orbiter in 2009. Most of these probes were outfitted with detectors with sensing capabilities across several sections of the electromagnetic spectrum, from X rays to infrared radiation. Because of the supposed water-ice deposits there, all spacefaring nations that expressed an interest in setting up human lunar operations concentrated on investigating craters near the South Pole.

NASA established the Constellation program in order to send astronauts back to the Moon. Initial flight operations were planned for late 2014 or early 2015, with the initial lunar mission perhaps coming within five years of those dates. If these missions were carried out, a new age of study of the Moon and its craters would begin.

APPLICATIONS

Because impact craters are instantaneous events, they are superb geologic time markers. Any material on which an impact crater and its ejecta are superposed is older than the crater; any material overlapping the crater or its ejecta is younger. Analysis of these relationships led to the development of the lunar stratigraphic column consisting of five systems. The pre-Nectarian system, comprising all lunar surface features formed prior to excavation of the Nectaris Basin, and succeeding systems defined by formation of the Imbrium Basin and the Eratosthenes and Copernicus craters.

Primary impact crater densities indicate the relative ages of different units on the lunar surface: the more craters, the older the surface. Craters employed in such studies are usually larger than 4 kilometers in diameter, and the densities are obtained by the extremely tedious task of simply counting them on a photograph. Crater density divided by the average crater-production rate gives the approximate absolute age of the surface units. In the case of the Moon, a calibration curve for production rate can be obtained by comparing the radiometric age of samples returned from the landing sites with crater-count statistics of those sites. These data indicate an exponential decline in crater production from about 4 billion years ago to the present. The details of crater production prior to about 4 billion years are the subject of debate, but, because most of the lunar surface postdates this period, the debate is of little relevance to age determinations.

In the middleground of this image from Apollo 15, the 40-kilometer-diameter crater Aristarchus can be seen to the left of the 35-kilometer-diameter Herodotus. (NASA)

Small craters (less than 3 kilometers in diameter) have also been used for dating purposes. These techniques are based on the fact that morphologies of small craters are modified in a consistent manner with time. One of these, called the D_L method, is based on the interior slope of the crater. As craters become progressively infilled, the length of the shadow cast by the rim decreases. For a given illumination angle, it is therefore possible to define the largest crater within an area that has reached a specified shadow limit. If a crater in another area is wider than the limit, the second area is older.

Predictable depth-diameter relationships of fresh impact craters allow the determination of some of the third-dimensional characteristics of lunar surface features. If a crater has been flooded by a younger lava flow, the extent of departure of that crater from the dimensions of a similarly sized fresh crater can be employed to determine the thickness of the lava (or other material). Effectiveness of this method is limited by the accuracy of the topographic data.

Material forming the lunar highlands has a different composition from that of the maria and results in pronounced spectroscopic differences. By analyzing spectroscopic signatures of ejecta blankets of craters superposed on the lunar maria, it is possible to determine if a crater excavated solely basaltic material or if it penetrated into the crust beneath. The depth-diameter relationship of the crater can then be employed to ascertain the mare thickness.

Mineralogical and geochemical analyses of returned lunar samples have played a large role in scientists' understanding of the physical processes involved during an impact event. Indirectly, lunar craters have also provided information on the deep lunar interior, because the impacts of both natural and human-made objects have generated seismic waves recorded by the Apollo seismic network.

For a long time claims were made by observers of the Moon that flashes of light came from certain areas of its surface, notably some large craters. Prior to the impact of the Ranger 8 spacecraft there, the crater Alphonsus, for example, was an area where some reported seeing flashes of light; their supposition was that Alphonsus had active volcanism. Most astronomers scoffed at the reports, and no serious observations of Alphonsus ever recorded undeniable light flashes. However, in late 2005 a coordinated search by astronomers at NASA began to record hundreds of small flashes of light representative of explosions on the order of a few hundred kilograms of TNT. The early supposition was that the flashes were the result of volcanic activity. Those recognized flashes were actually from impacts of fragments of the extinct comet 2003 EH1, the progenitor of the Quadrantid meteor shower. This observation indicated ongoing alteration of the lunar surface, a process that had long been attributed to the production of the lunar regolith but was never before seen occurring routinely. The influx of material creating microcraters hinted at a threat to astronauts returning to the Moon to stay and establish permanently occupied bases for research.

CONTEXT

The heavily cratered surface of the Moon provides a scenario for what was also taking place on Earth at a time for which scientists have no geologic record. They have learned that, as one goes farther back in time, the number of objects that hit the Moon increases exponentially until around 4 billion years ago. The study of lunar craters has influenced the conceptualization that meteorite impacts may have played a role in terrestrial mass extinctions and thereby the evolution of life. It has been suggested that, because the Moon contains so many craters, a large number of age determinations of the impact events could provide information concerning a hypothesized correlation between impact bombardment and cyclic extinctions.

Analyses of the sizes and distribution of volcanic craters have provided data concerning the internal thermal evolution of the Moon and the stress distributions within the upper lunar crust. In addition, study of the morphologies of impact craters allows scientists to determine the effects of impacts within a gravitational field one-sixth that of the Earth and on a body with no atmosphere. Much of the work by Ralph Baldwin in formulating the characteristics of impact craters was based on data from small terrestrial human-made explosions. Scientists are now able to predict fairly accurately the consequences of very large explosions on land, into water, or in space as a result of lunar impact crater studies. Much of Baldwin's original work remains viable even decades after the first lunar landings.

It has been suggested that the permanently shadowed floors of some near-polar craters may be reservoirs for trapped volatiles such as water. Such resources, if present, would play an important role in the location of a crewed lunar base. Furthermore, impact craters have played the major role in forming the lunar regolith. This loosely aggregated material could be mined with limited mechanical processing. Conversely, the myriad craters on the lunar surface pose a major hazard for safe surface travel and will probably result in unavoidable detours for initial crewed expeditions. Both structures and persons would have to be protected against the small impacting bodies that rain onto the lunar surface.

James L. Whitford-Stark

FURTHER READING

Baldwin, R. A. *The Measure of the Moon*. Chicago: University of Chicago Press, 1963. The first nine chapters outline the characteristics of both lunar and terrestrial, man-made, and natural craters. Written in pre-metric-system times, this classic text requires mathematical (though simple) conversions. Suitable for advanced high school and college students.

Beattie, Donald A. *Taking Science to the Moon: Lunar Experiments and the Apollo Program*. Baltimore: Johns Hopkins University Press, 2003. Explains the science gleaned from the Apollo lunar landings, including the Apollo Lunar Surface Science Experiment Packages (ALSEPs) and their results.

Chaikin, Andrew. *A Man on the Moon: The Voyages of the Apollo Astronauts*. New York: Penguin, 2007. Reissue of one of the most engaging accounts of the Apollo program and the exploration of the Moon. Critically acclaimed. For all ages.

Consolmagno, Guy, and Martha Schaefer. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994. A text accessible at the college level for science and nonscience readers alike. Presents subjects at low-level mathematics and also involves integral calculus where required. Demonstrates how the area of planetary science progresses by questioning previous understanding in the light of new observations.

Grego, Peter. *The Moon and How to Observe It*. New York: Kindle Books, 2005. For the amateur observer just starting out, this is the guide that will make one's initial steps at backyard astronomy more enjoyable. Information of use to the more skilled observer too. Primarily uses the Moon as its target in explaining how to make and record observations with telescopes and camera systems, both analog and digital.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Strong on Earth-Moon geology. Takes a comparative planetology approach instead of presenting individual chapters on each planet in the solar system.

Jefferis, David. *Return to the Moon*. New York: Crabtree Children's Books, 2007. An explanation of contemporary plans to return American astronauts to the Moon as part of Constellation program operations with the Orion Crew Exploration Vehicle and Altair landing craft. Aimed at a younger audience.

Melosh, H. J. *Impact Cratering: A Geologic Process*. New York: Oxford University Press, 1996. A good synthesis of research information related to the impact process. Mathematics separated out in text and in two appendixes. Aimed at graduate student and research audiences.

Schmitt, Harrison J. *Return to the Moon: Exploration, Enterprise, and Energy in the Human Settlement of Space*. New York: Copernicus Books, 2006. A scientific and economic plan for prolonged exploration and exploitation of lunar resources written by the only geologist to land on the Moon and perform field geology and collect documented samples for return to Earth. Talks about helium-3 mining in order to supply future energy production systems on Earth and thereby finance and systematically expand lunar settlements and research posts.

Schrunk, David, Burton Sharpe, Bonnie Li Cooper, and Madhu Thangavelu. *The Moon: Resources, Future Development and Settlement*. New York: Springer Praxis, 2007. Examines what is needed to develop human settlements on the lunar surface beginning with renewed robotic exploration. Explains the benefits of a robust spacefaring civilization.

Spudis, Paul D. *The Geology of Multi-ring Impact Basins: The Moon and Other Planets*. New York: Cambridge University Press, 1993. A comprehensive geological study of large impact craters formed on the planets and their satellites. Extensively illustrated with photography and diagrams. Suitable for both the lay reader and scientific audiences.

Wilhelms, D. E. *The Geologic History of the Moon*. U.S. Geological Survey Professional Paper 1348. Washington, D.C.: Government Printing Office, 1987. Full of photographs of lunar craters supplemented with an outstanding text based on more than twenty years of lunar geological mapping by the author. Research-level book but understandable to the dedicated high school student. A must-have book for those interested in the Moon.

See also: Earth-Moon Relations; Impact Cratering; Lunar History; Lunar Interior; Lunar Maria; Lunar Regolith Samples; Lunar Rocks; Lunar Surface Experiments; Meteoroids from the Moon and Mars; Solar System: Origins.

Lunar History

Category: Natural Planetary Satellites

The heavily cratered lunar surface has slowly evolved over the past 4.6 billion years primarily because of impact events. Observations made by astronomers with ground-based telescopes, studies carried out a result of the Apollo program, and direct studies of Moon rocks have revealed much about this formation process.

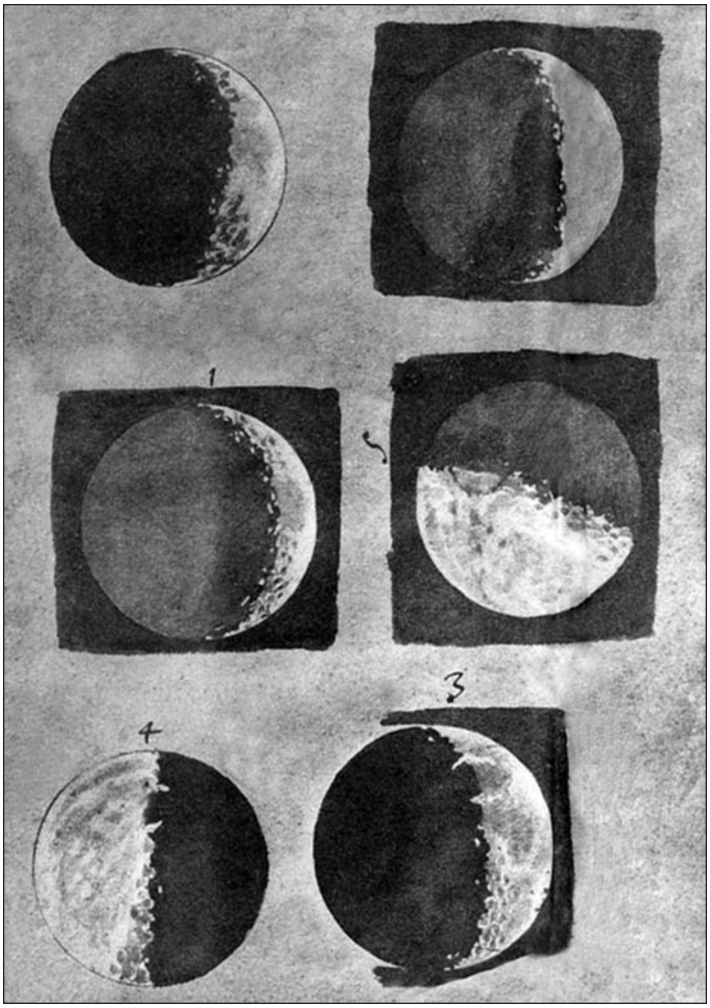
OVERVIEW

The Moon's surface has evolved into its present state as a result of meteoritic impact over the eons since the solar system formed. In contrast, Earth's surface has been molded and shaped primarily as a result of geologic activity brought on by heat transfer from its molten core. One can argue that these very different surfaces—the Earth's constantly changing, eroding surface and the Moon's relatively tranquil, slowly changing surface—are the result of each body's total mass. Generally, low-mass bodies acquire surfaces with meager geological activity, whereas more massive bodies continuously undergo a considerable amount of geologic change. There are thought to have been six states in the evolution of the Moon: the origin of the Moon, the separation (or differentiation) of the Moon's crust, the first age of igneous activity, the great bombardment period, the second age of igneous activity, and the quiescent period.

The origin of the Moon is the stage that is least understood. Astronomers, physicists, geologists, and mathematicians have struggled and debated over the origin of the Moon for at least two centuries. The original theory of the Moon's formation was that it was a body accreted elsewhere within the solar system and came close enough to the Earth to be captured by its gravitational field. However, such a hypothesis proved difficult to model mathematically and was eventually proven to be impossible in a two-body system with masses such as those involved in the Earth-Moon system.

Another hypothesis was that the Earth and the Moon formed about the same time, shortly after the proto-Sun (the very early stage of the Sun, which contracted by gravity out of interstellar material) some 4.6 billion years ago. The Earth and the Moon formed relatively near each other (at a distance that is small compared to the distance between planets). This theory lost favor, however—especially after determination of the chemical composition of Moon rocks. Lunar samples were devoid of many common elements in Earth rocks. Because the Moon lacks water, the number of lunar minerals relative to those found on Earth is quite small.

A modern hypothesis of the lunar origin is called the large impact hypothesis. It gained favor among scientists and then became subject to intense investigation and scrutiny. The idea is that a Mars-sized body struck the early Earth. Such a catastrophic collision would have resulted in the ejection of a large cloud of debris from the obliteration of the Mars-sized object and severe damage to the early Earth. Subsequently, the material would have condensed into a disk of debris, some of which would have fallen back to Earth, completing the formation of the Earth. In time the rest would have accreted into the Moon. This theory has many features that are consistent with the knowledge that the Earth and Moon formed near each other and explains the similarity of the Earth's



The astronomer Galileo published these sketches of his lunar observations in Sidereus Nuncius *in 1610.* (The Galileo Project, Rice University)

surfaces of the Moon and other planets. These meteoritic showers were especially frequent in the early solar system. Since the outer layers of the Moon were molten, the lighter minerals tended to float and the heavier ones tended to sink. This means that heavy igneous material, such as the mineral basalt, would sink, whereas lighter igneous material, such as the mineral anorthosite, would float. The Moon's crust differentiated into an igneous shell with an inner zone of basalt and an outer zone of the lower-density igneous rock anorthosite. Nearly all specimens collected during the Apollo program can be put into three categories: mare basalt, found in the maria (the basins, from the Latin for "seas"); KREEP-norite, with an unusually high content of potassium (K), rare Earth elements (REE), and phosphorus (P); and the anorthosite group. The latter two are found in the lunar highlands.

As the Moon's surface solidified, several pockets or subsurface pools of KREEP-norite remained molten because of slight differences in chemical composition within the crustal rock. This episode of igneous activity is the third stage of lunar evolution. Certain impurities would form crystals with lower melting temperatures and thus remain liquid. Some of the darker shades of pale-yellow-rock highland material, surrounding the maria, are KREEP. Coincidentally, meteorites struck above some of these liquid pockets and fractured the solid rock. The cracks and fractures formed would occasionally lead to one of these pools, and the release in pressure would cause the liquid to "creep out" onto the surface.

The fourth stage, the great bombardment period, overlaps all the other stages. Its placement in the list of the Moon's evolutionary stages is based on the concept that this age is easily identified by the largest craters still visible today. Continuously during the first 500 million years and the first four stages, impact-

crust to the composition of the Moon. It also explains why Earth's moon is so much larger in proportion to its parent planet than any other satellite in the solar system.

The second stage of evolution, the differentiation of the crust, occurred during the early stages of origin, when the Moon was still molten in its outer layers. This condition occurred from the heat of formation and was prolonged by constant bombardment of meteorites striking the Moon's outer layers. In this early stage of the solar system, much debris was still in the form of a complex set of rings around the Sun. Chunks of the debris, composed of meteoroids, fell onto the ing bodies of all sizes struck the Moon. The largest formed multirimmed basins, hundreds of kilometers in diameter. The meteoroids responsible may have been huge bodies 150 to 200 kilometers in diameter. Many of the largest craters observed in the highlands date back to impacts during this period. This is also true of many of the large basins, which in turn are often accompanied by the gray-blue colored maria.

As the result of a high-speed impact, both the meteoroid and the underlying surface material of the Moon were subject to rapid, intense heating. All the kinetic energy of the projectile's motion was translated into other forms, namely heat, shock waves, and energy of excavation of lunar material, since the motion of the impacting body was halted abruptly. In a fraction of a second, outer regions of the meteoroid and some surface material were vaporized. The resulting high temperature caused dense gas to explode violently. This activity resulted in a crater surrounded by a rim, composed mostly of the material that was ejected from the resulting hole. Beyond the rim, arranged in a sunburst or spray pattern, was the ejecta blanket, complete with innumerable small craters (called secondaries) of a wide range of sizes.

Since the mass of the Moon is low compared to that of Earth, the Moon lost most of its internal heat of formation and the heat provided by radioactive decay in a short time, perhaps less than 800 million years after its origin. Earth, being eighty-one times more massive than the Moon, has retained much of its heat, as a result of radioactive decay in its interior. Although the Moon may still possess a molten core, it does not transfer a substantial heat flow to the surface, as it once did. Heat-flow experiments performed by Apollo astronauts on the Moon's surface did register an outflow of heat, but those experiments were rather rudimentary. More sophisticated and far-ranging detectors are needed to determine the Moon's internal thermal state more precisely.

Between 3.8 and 3.1 billion years ago—during the fifth stage of lunary history, the second great igneous period—only the lower basaltic strata of the separated crust were heated. The largest impacts, slightly after and during this stage, formed large fractures that in some cases reached down to the lower basaltic layers. This allowed molten basalt to flow as lava to the surface and fill in the low-lying areas in and around the basins. This accounts for the gray-blue basaltic maria. All these maria have generally circular shapes, with one exception, the Ocean of Storms. This pancake shape is what one would expect from flows into large, almost circular craters. Since basalt has a higher density than anorthosite, these pancaked lava flows have been detected by the effect their gravity has on satellites orbiting the Moon. As a result of their localized pull of gravity, they have been named mascons (mass concentrations). The flow of basalt onto the lunar surface apparently came in stages, since astrogeologists can identify younger flows atop older flows. The age of the flows can be identified by comparing the frequency of impact craters found in the flows and also the wrinkle structures in the maria that identify the border of a flow.

Finally, after the first one billion years, the igneous activity ceased and impacts gradually became less frequent. This final stage, the quiescent period, lasted more than three billion years to the present time.

METHODS OF STUDY

Almost all the features on the Moon's surface can be shown to be caused by impacts—directly, as in the case of a crater, or indirectly, as in the case of KREEP or a maria. There are no mountains on the Moon resulting from uplifting. All the mountainlike features are partial or quite complete rims of basins or craters. The partial rims have been obliterated by more recent impacts. Another mountainlike feature often seen at the center of a well-preserved or relatively fresh (young) crater results from the focus of a shock wave reflected and focused by the rim formation during impact back toward the impact center. Arrival of the reflected shock wave apparently heaves the center of the crater upward.

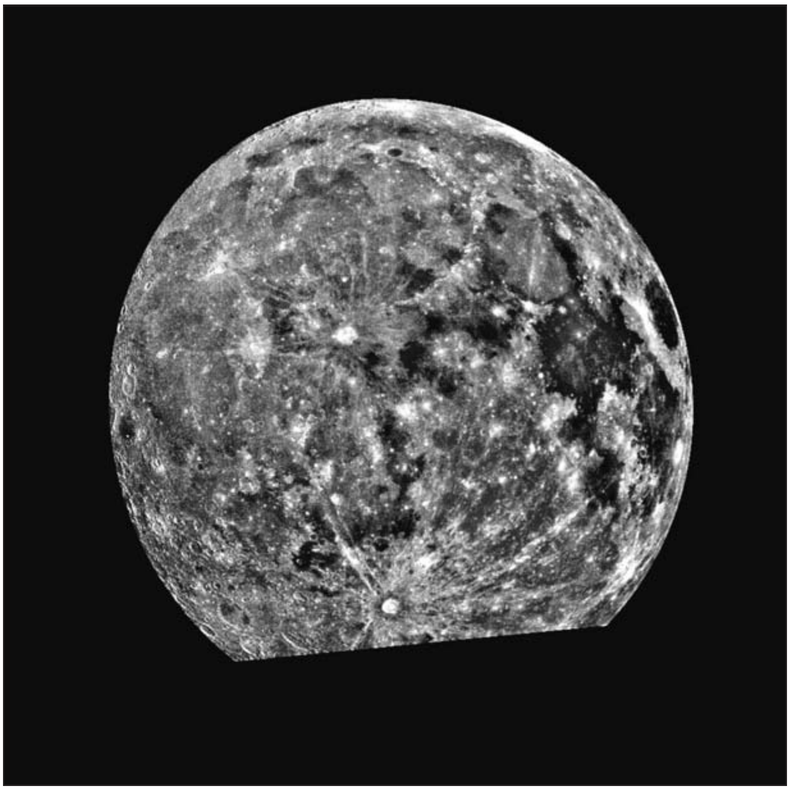
The surface of the Moon today is covered with a finely pulverized dust called regolith. Regolith varies in depth from 5 to 10 meters in young maria (3.1 billion years old) to perhaps 20 to 25 meters or more in the highlands (more than 4.2 billion years old). Regolith has been formed by the constant churning and mixing from particles impacting through the ages, ranging in size from micrometeorites to meteoroids large enough to form the largest craters. Crater rims, having both a sharp (young) and a rounded, smoothed (old) appearance, are evidence of this same kind of slow erosive process caused by impacts.

Close microscopic inspection of the regolith reveals evidence of impacts and the Moon's evolutionary history. The regolith consists of tiny pieces of anorthosite, anorthositic breccia, basalt, basaltic breccia, a variety of glasslike, irregularly shaped particles, and small, spherically shaped glass beads. The glass is believed to be produced by the heat and shock of impact, resulting in a fused metamorphic, glasslike rock. Spherical beads result from liquid drops thrown out during the splash of ejection shortly after impact. The drops freeze or solidify before falling back to the surface.

In the early 1970's, Apollo astronauts conducted field geology exercises on the Moon's surface and placed experiments there to study moonquakes in the hope of getting information about the lunar interior. Moon rocks brought back to the Earth played a major role, if not the most important role, in advancing understanding of the evolution of the lunar surface. Fragmented pieces of lunar rock provide many clues to lunar history.

Geologists can easily identify several distinct rock types in such samples. Some of the gray pieces are the well-known igneous rock basalt. Some of the very white pieces are brecciated anorthosite; breccia is rock that has been fragmented and welded together with immense heat. The glasslike substances and those resembling glass beads provide an understanding of the evolution of the Moon's surface as brought about by eons of meteoritic bombardment. Apollo studies, together with previous space-based and earthbound telescopic studies, have allowed scientists to construct an evolutionary history of the lunar surface.

All lunar samples are geologically processed. Their composition was established by igneous processes inside the Moon. No primitive or primordial lunar material (material that existed on the Moon's surface soon after the formation of the Moon) is likely to have survived the turbulent early history of the Moon. Primordial material from which the Moon was formed should be about 4.6 billion years old. The Earth and Moon, moreover, formed in the same general region. This is indicated by the specific abundances of certain types of atoms, especially those of oxygen. Abundances are good indicators of position from the Sun in the inner solar system. Lunar



Even in black and white, this false-color image—compiled from 15 different images of the Moon from the Galileo mission in late 1992—highlights the heavily pockmarked surface; false color was used to assist planetary geologists in identifying the types of rock and minerals that dominate different regions. (NASA/JPL) and crustal terrestrial rocks have very similar abundances. The overall density of the Moon is similar to the density of the Earth's crust; for this reason, scientists in the past speculated that the Moon came from the Earth.

Neither humans nor robots have visited the lunar surface and returned samples to Earth since Luna 24 in 1976. The Ballistic Defense Organization used the Moon as a test bed in early 1994 for an evaluation of next-generation sensors flown on the Clementine spacecraft. This small probe, which lasted only 115 days, carried seven different instruments capable of examining the lunar surface across much of the electromagnetic spectrum. In coordination with the National Aeronautics and Space Administration (NASA) a mineralogical map of much of the Moon was obtained. Surprisingly, there was a signature strongly suggesting that the Moon possessed a significant amount of water ice near the poles. Clementine was followed four years later by NASA's Lunar Prospector, which expanded upon the Clementine science. Over the course of nearly eighteen months, Lunar Prospector orbited the Moon and assembled data that provided a detailed map of surface composition and identified lunar resources that might be used to support lunar bases.

In the aftermath of the *Columbia* accident (February 1, 2003), the Bush administration directed NASA to proceed with a next-generation crewed spacecraft that could again take astronauts beyond low-Earth orbit, expand the human presence to return to the Moon to stay, and eventually fly to Mars and beyond. Rather than setting up a time-limited crash program like Apollo, the Bush administration established the Space Vision for Exploration, which was open-ended. NASA responded to the challenge with the Constellation program, which included development of new boosters to send an Orion Crew Exploration Vehicle to the Moon with a much larger and more versatile Altair Lunar Lander in order to establish a base. The precise site of the base remained under evaluation, but its general vicinity would be near the Moon's south pole to make use of the suspected water ice deposits there and to also be able to take advantage of permanent solar irradiation for power generation. Plans called for a return to the Moon with Constellation assets by 2019, the fiftieth anniversary of the first crewed lunar landing (Apollo 11 in July, 1969). In addition to setting up infrastructure at the Moon's south pole, Constellation was to be designed to support exploration about the lunar surface far from the base itself to attempt to answer many of the questions about the Moon's history and geology that Apollo did not completely answer.

The United States was not the only nation at this point with interest in the Moon. After much indifference, international interest in understanding the Moon's history and making determinations about potential uses of its resources began to increase dramatically in the early twenty-first century. China began sending taikonauts (the Chinese word for astronauts) into space in Shenzhou spacecraft in late 2003. The Chinese crewed space program was designed to incrementally advance to lunar missions. Indeed, it is possible that Chinese taikonauts will reach the Moon before NASA astronauts return. NASA's next robotic spacecraft dedicated to lunar science is the Lunar Reconnaissance Orbiter (LRO). LRO joins probes sent to the Moon by the European Space Agency (SMART 1 spacecraft), the Japanese Aerospace Exploration Agency (Kaguya spacecraft), and the Indian Space Agency (Chandraayan-1 spacecraft).

CONTEXT

Understanding the history of the lunar surface has helped scientists to discern the events and processes that led to the origin of the solar system, the planets in general, and the Earth-Moon system in particular. It is important to understand the history of the Moon in that it provides information about the history of the Earth. Because of the Moon's ancient surface, questions that cannot be answered by examining the youthful terrestrial surface have been answered by lunar studies.

Detailed studies of lunar rocks have led to a much deeper understanding of the probable chemical composition of the solar nebula and the proto-Sun. This is important not only because it helps explain the process whereby the Sun and planets formed but also because it provides specific clues and observations that increase scientists' understanding of star formation in general. Astronomers now believe that planetary formation is a natural consequence of star formation.

Anomalous abundances of particular chemical elements found in both lunar and terrestrial rock have led directly to speculations about the origin of the solar system. Some chemicals in the rare Earth group strongly indicate that a supernova explosion may have contaminated the gas and dust from which the solar nebula formed with chemical debris from deep within the exploding star. Some astronomers suggest that the shock of the supernova blast wave actually triggered the contraction of the solar nebula into the formation of the Sun and planets.

Understanding the lunar surface also advances the goal of its future exploitation. The Moon might someday be used as a space station, an astronomical observatory, or a space colony. The Moon would be an excellent source of minerals, since virtually all lunar rocks are rich in the metal titanium and would supply much more of this metal than could be mined on Earth from even the richest deposits. To these ends, NASA is developing the Constellation program, which includes returning humans to the Moon. The goal of Constellation is to expand the human presence beyond low-Earth orbit. The first major step in that lofty, open-ended adventure includes the development of a permanently staffed lunar base, one from which exploration and science can be conducted, all the while making use of lunar resources to meet the needs of the scientists and astronauts living and working on the Moon.

James C. LoPresto

FURTHER READING

Abell, George O., David Morrison, and Sidney C. Wolff. *Exploration of the Universe*. 7th ed. Philadelphia: Saunders College Publishing, 1995. This elementary astronomy textbook is considered by many astronomers to be one of the finest available. It is written in a traditional style, and the chapter on the Moon includes several good photographs, images, and diagrams.

Baldwin, R. A. *The Measure of the Moon*. Chicago: University of Chicago Press, 1963. The first nine chapters outline the characteristics of both lunar and terrestrial, human-made, and natural craters. Written in pre-metric-system times, this classic text requires mathematical (though simple) conversions. Suitable for advanced high school and college students.

Beattie, Donald A. *Taking Science to the Moon: Lunar Experiments and the Apollo Program*. Baltimore: Johns Hopkins University Press, 2003. Explains the science gleaned from the Apollo lunar landings, including the Apollo Lunar Surface Science Experiment Packages (ALSEPs) and their results.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory text covering the field of astronomy. Contains descriptions of astrophysical questions and their relationships. Informative.

Grego, Peter. *The Moon and How to Observe It*. New York: Kindle Books, 2005. For the beginner or amateur observer, this is the guide that will make one's initial steps at backyard astronomy more enjoyable. Primarily uses the Moon as its target in explaining how to make and record observations with telescopes and camera systems, both analog and digital. Of use to more skilled observers, too.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Explains all current theories for the formation of the Moon.

Kosofsky, L. J., and Farouk El-Baz. *The Moon as Viewed by the Lunar Orbiter*. NASA SP-200. Washington, D.C.: Government Printing Office, 1970. A fine collection of early lunar images taken from the Lunar Orbiter satellite series. They consist of both overall global images and detailed high-resolution images of particular areas on the Moon.

Wilhelms, Don E. *The Geologic History of the Moon*. Professional Paper 1348. Denver, Colo.: U.S. Geological Survey, Books and Open-File Reports Section, 1987. A detailed publication on lunar history with a strong emphasis on what has been learned as a result of the space program.

______. *To a Rocky Moon: A Geologist's History of Lunar Exploration*. Phoenix: University of Arizona Press, 1994. Numerous books portray the Apollo astronauts, who successfully completed six lunar landings, with heroic prose. Others describe the science learned from Apollo investigations and returned Moon rocks and soil samples. This book is both a personal account and a geology-driven explanation for the way individual missions were planned and carried out for maximum science return. Includes descriptions of real-time evaluations of field geology aspects on each Apollo mission.

See also: Earth-Moon Relations; Impact Cratering; Lunar Craters; Lunar Interior; Lunar Maria; Lunar Regolith Samples; Lunar Rocks; Lunar Surface Experiments; Meteoroids from the Moon and Mars; Planetary Satellites; Solar System: Origins.

Lunar Interior

Category: Natural Planetary Satellites

Using fundamental knowledge of the Earth's interior along with data returned by missions to the Moon, scientists have been able to extrapolate theories about the lunar interior.

OVERVIEW

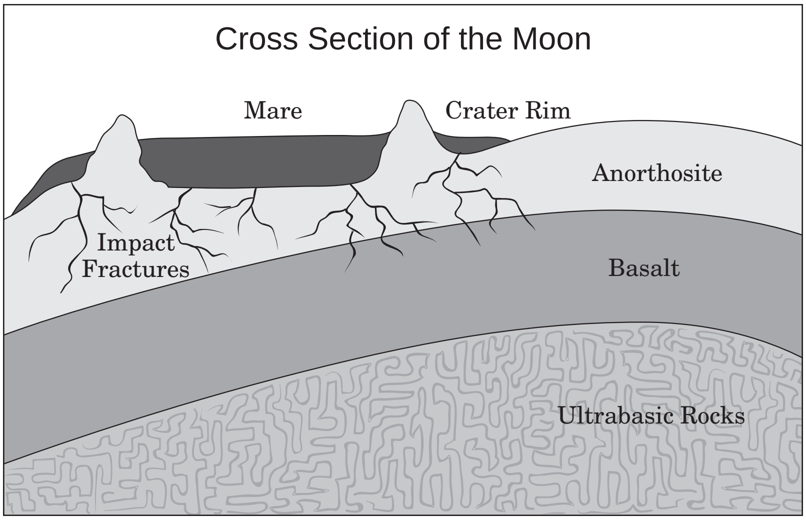
The contrasting light-colored highlands and darker maria regions of the Moon were very apparent to even the earliest observers. Early telescopic observations revealed the highlands to be rough, cratered, and mountainous, as compared to the smoother and less cratered maria. Galileo's first impression of the lunar surface clearly drew a comparison to the land and sea regions of the Earth. Later scientists, with only telescopic observations to guide them, assumed that the Earth and Moon probably had similar origins and should be quite alike in most respects. This assumption did not last long.

When scientists made their first calculations of the densities of the Earth and Moon, they discovered an interesting fact. The Earth has a density of 5.5 compared to 3.34 for the Moon. This seemed strange for two bodies found so close to each other in space. Later theories attributed the difference in densities to the Earth's having a much higher percentage of metal in its overall chemical composition. With the higher metal content, the Earth would naturally have a greater mass and a higher density. Two distinctly different sets of characteristics would soon define each object, for it is mass that creates a planet's internal pressure and temperature conditions. These two factors, in turn, determine whether or not a metal or silicate mineral will remain a solid or become a molten liquid. Once a molten liquid is produced, the process of density separation can take place and produce a distinct core.

Based on a density comparison between the Earth and the Moon, it is believed that the Moon does not have a well-defined core like that of the Earth. Experiments conducted on the Apollo 15 and 17 missions did indicate that the Moon's heat-flow rate is about half that of the Earth. Although the lunar interior is hot, it is not sufficiently hot to produce a density separation of materials comparable to that which occurred in the Earth. The higher iron content of the lunar crust also tends to support a planetary body that is not well differentiated like the Earth.

The current model for the nature and chemical composition of the Moon's interior has been primarily derived from lunar rock samples and the seismic experiments left on the Moon by the Apollo astronauts. This information, combined with data from both crewed and uncrewed orbital missions, has given scientists a much clearer idea of what constitutes the lunar interior. Originally, based on the relatively well-established models developed for the Earth, it was assumed that the Moon's interior should have experienced a similar history. As Apollo data started to pour in, it quickly became apparent that this was not going to be the case. In subsequent years, differences between the interiors of the Earth and Moon have become even more controversial.

The interior of the Earth has been divided into a crust, mantle, and core. This division is based on the determination of the densities of various rocks and metals at specific depths. This division does not represent the primordial Earth. One theory of the origin of the Earth suggests that the Earth formed as a result of the accretion of innumerable cold, solid bits of rock and metal into a relatively cold, chemically homogeneous body. Shortly after the Earth's formation, its internal temperature began to rise due to the decay of radioactive isotopes, the heat from accretion, and the mass of the Earth itself. At this point, its interior temperature was well below the melting point of most metals and silicate minerals. Gradually temperatures reached a point where melting occurred, producing a molten liquid. Within this molten magma, denser metals sank to the center of mass forming the liquid and solid metallic cores. As a result, lower-density silicate minerals were displaced and moved upward to form the crust.



Even within Earth's crust itself, rock materials of two distinctly different densities would separate into the lower-density continental crust and the higher-density oceanic crust. Sandwiched between the crust and core would be the mantle, a large region of silicate minerals with variable densities appropriate to specific depths. This entire process is referred to as density separation, and it is believed to have taken place very early during the initial stages of the Earth's formation. Through this separation of molten metals and nonmetals, the generation of an electric current and magnetic field became possible.

If this process occurred for Earth, then could a similar process have taken place on the Moon? Perhaps the best indicator of the composition and structure of the lunar interior came from the Apollo Lunar Science Experiment Packages (ALSEPs), which included a seismometer to record "moonquakes." On Earth, scientists use seismic waves to calculate the density and predict the composition of materials at various depths. The way seismic waves pass through various materials gives clues to their chemical and physical properties. The same should be true for the Moon and thus provide a detailed picture of the lunar interior. Apollo results indicate that the internal structure of the Moon is very different from that of the Earth. Both the lunar crust and mantle are much thicker relative to the Earth's and show no evidence of plate tectonics. Seismic evidence does indicate that the upper portion of the lunar crust has been shattered by countless meteoroid impacts; with depth, this crust gradually progresses into a solid rock layer termed the lithosphere. Beneath this first 1,000 kilometers (620 miles) lies the asthenosphere, a region where seismic waves have indicated the presence of a liquid or partially liquid environment, which may include the core. The core, if it truly exists, may consist of an iron sulfide mixture rather than a "pure" nickel-iron alloy. The presence of a significant amount of iron sulfide minerals could lower the melting temperature required to produce a liquid phase and hence the conditions suitable for the formation of a core. The specific chemical composition of the lunar interior, along with the existing temperature and pressure conditions, are the defining factors as to whether or not the Moon has a core similar to Earth's. Existing data suggest that the Moon does not have sufficient internal heat to produce a distinct metallic core. The fact that iron-bearing lunar materials are not magnetic also points to the apparent lack of a magnetic field, which is usually attributed to the presence of a liquid-solid metallic core.

KNOWLEDGE GAINED

The arrival of the "space age" gave scientists the opportunity to answer many of their questions about the origin and evolution of the planets, because new technology allowed them to send satellites and probes to explore other worlds at close range. Because of its close proximity to Earth, the Moon was a logical first choice for exploration. Several hundred years of telescopic studies had not provided many answers; to unravel the Moon's secrets, both crewed and uncrewed spacecraft would need to go to the Moon and collect data.

Beginning in the early 1960's, when President John F. Kennedy challenged the American people to go to the Moon, the National Aeronautics and Space Administration (NASA) developed a series of lunar exploration programs: the Ranger, Lunar Orbiter, Surveyor, and Apollo missions achieved that goal. Ranger provided the first close-up look at the Moon. Lunar Orbiter provided the reconnaissance images to select the landing sites for Surveyor and later the Apollo missions. In addition to proving that a lunar soft landing was possible, Surveyor gave science its first look at the lunar soil and surface conditions. The six Apollo missions not only returned more than 380 kilograms of lunar rock and soil but also left experiments on the lunar surface to study the Moon's interior. Later missions, such as Clementine and the Lunar Prospector, surveyed the lunar surface for mineral deposits and searched for the presence of water ice. The Russian Luna program also added to our overall knowledge of the Moon through the use of a robotic rover and sample-return missions.

In 1998 astronauts aboard space shuttle Discovery *took this picture of the Moon from Earth orbit, as seen to the right of the Atlantic Ocean through the shuttle's aft windows.* (NASA)

Knowledge seems to flow back and forth between the Earth and Moon. The basic geological principles geologists have learned on Earth have been applied to lunar features, yet many of the discoveries made on the Moon have caused geologists to rethink their original theories. In the context of a wealth of lunar data, scientists now see the Moon as a world seemingly similar to Earth yet markedly different. The early processes that created a distinct crust-mantle-core structure for the Earth and produced a strong magnetic field never reached completion on the Moon. The Earth was able to retain its high internal temperature and remain fluid at specific depths, while the Moon apparently did not and remains only a partially differentiated body. Further studies of the Moon are certainly needed before science can provide a definitive understanding of the Moon's physical makeup and interior structure. Future lunar missions may answer many of the remaining questions concerning the lunar interior as well as giving a better understanding of lunar surface materials and the giant impact processes that have shaped lunar history.

While lunar geologists eagerly await a return to the Moon, analysis of samples returned by the Apollo astronauts continued decades after the last lunar landing, Apollo 17 in December, 1972. A rock returned on that mission after being collected by Dr. Harrison Schmitt, the only geologist to land on the Moon during Apollo, turned out to be the oldest sample collected that had not been subjected to intense shocks from major bombardments of the Moon occurring after the rock's formation. This rock possessed a remnant magnetism dating back to beyond 4.2 billion years, indicating that in its early history the Moon had a liquid core that produced magnetism by a dynamo effect, as Earth still does. That lunar field appears to be about one-fiftieth that of the Earth, a result that models of lunar core dynamics also predict. This finding adds evidence to the theory that the Moon did not form cold but out of a collision between a Mars-sized object and the early Earth.

CONTEXT

As far back as 400 B.C.E., ancient Greek philosophers wondered about what lies beneath the Earth's surface. They envisioned a dark, hot, sulfurous underworld populated by demons and the spirits of the dead. This dismal picture was based on a certain amount of truth. The ancients were familiar with volcanoes and sulfurous hot springs and easily made the connection to the underworld.

Today modern science certainly has an advantage over the Greeks. Scientists now have the ability to study the Earth's interior by means of seismic studies, deep drill holes, and analyses of deep-seated rocks brought up during violent volcanic eruptions. With data like these, it is possible to develop computer models that give a very accurate picture of what lies deep beneath the Earth's crust. Once we have a basic idea of what makes our planet work, natural human curiosity causes us to wonder if these same geological processes are common to other worlds. For more than fifty years, humankind has extended its curiosity to our Moon and found a world that is both familiar and different from Earth. Certainly, future lunar missions will not only expand our knowledge of the Moon's interior but also pave the way for a more comprehensive understanding of the terrestrial planets.

Paul P. Sipiera

FURTHER READING

Canup, R. M., and K. Righter, eds. *Origin of the Earth and Moon*. Tucson: University of Arizona Press, 2000. A compilation of twenty-nine scientific papers dealing with terrestrial planetary formation with emphasis on the Earth-Moon system. Excellent reference source for undergraduate and graduate students.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A comprehensive examination of our solar system from the planetary geologist's perspective. Suitable for advanced high school students and undergraduates.

Hubbard, William B. *Planetary Interiors*. New York: Van Nostrand Reinhold, 1984. A classic work that addresses the fundamentals of planetary interiors with detailed coverage of the Earth and Moon. Appropriate for high school and college-level readers.

Lang, Kenneth R. *The Cambridge Guide to the Solar System*. Cambridge, England: Cambridge University Press, 2003. A concise yet comprehensive reference for all members of the solar system, with detailed information on the lunar interior. Suitable for a wide range of readers.

Mohit, P. Sundas. "The Two-Faced Moon." *American Scientist* 96, no. 3 (May/June, 2008): 210-217. Addresses the intriguing issue of the Moon's variable crustal thickness with many references to the nature of the lunar interior. Suitable for a general audience.

Mutch, Thomas A. *Geology of the Moon*. Rev. ed. Princeton, N.J.: Princeton University Press, 1972. A classic work written by one of the leading planetary geologists of the Apollo era. Despite its publication date, this work still serves as an excellent starting point for an understanding of Earth's moon. Suitable for the general reader.

Taylor, Stuart Ross. *Planetary Science: A Lunar Perspective*. Houston, Tex.: Lunar and Planetary Institute, 1982. An often-referenced source in the current literature, despite its age. Provides the reader with fundamental information that aids understanding of later developments in lunar science.

See also: Earth-Moon Relations; Impact Cratering; Lunar Craters; Lunar History; Lunar Maria; Lunar Regolith Samples; Lunar Rocks; Lunar Surface Experiments; Meteoroids from the Moon and Mars; Planetary Satellites.

Lunar Maria

Category: Natural Planetary Satellites

The Moon's maria (literally "seas") are low-elevation areas of the lunar surface, in contrast to the lunar highlands. Interest in lunar maria achieved international prominence in 1969, when the National Aeronautics and Space Administration's (NASA's) Apollo 11 astronaut Neil Armstrong set foot upon the surface of the Moon in the Sea of Tranquility. There Armstrong and fellow astronaut Edwin E. Aldrin collected the first samples of the surface of another natural body in space.

OVERVIEW

Mare (plural, *maria*) is the Latin for "sea"; the term reflects seventeenth century ideas about the nature of the lunar surface. Similarly, there are other places on the lunar surface named for bodies of water: the Oceanus Procellarum (Ocean of Storms), some areas bearing the name *lacus* (lake), others called *sinus* (bay), and still others called *palus* (marsh). Returned lunar samples have established not only that there is no surface water on the Moon but also that the Moon as a whole has so little water that there are no hydrated minerals in the rocks. In reality, the maria are areas of basaltic lava flows. These lavas occupy approximately 16 percent of the lunar surface, with 80 percent of those located in the equatorial area on the side of the Moon that is always turned toward the Earth—a total of more than 6 million square kilometers. Although extensive, mare basalts probably represent less than 1 percent of total lunar crust by volume.

Volcanoes form some of the highest features on the Earth, but the majority of the lunar lavas are situated in low regions. These lows were excavated by meteorite impacts that left craters varying in size from microscopic to 2,500 kilometers in diameter. The largest of these craters, called basins, were later filled by lavas to form the maria. Smaller craters, particularly those immediately surrounding the nearside maria, also became the sites of lava eruptions. Estimates of the thicknesses of the mare lavas range from less than 100 meters in the shallowly flooded craters to perhaps 10 kilometers in the larger basins. Depth of basalt fill appears to be related to the age of the impact-basin-forming events, the deeper fill being found in the younger basins. Nearly thirteen hundred separate eruption locations have been identified. This number does not include those buried by their own or younger erupted materials. The true number of eruption sites may be closer to thirty thousand.

Returned lunar samples have radiometric ages that fall within a range of 3 to 4 billion years. The presence of basaltic fragments in breccias produced by the large-basin-producing impacts dating from about 4 billion years, along with dark materials excavated by impacts from beneath a younger impact depositional cover, indicates that the age of volcanism extends further back in time than the ages of the returned samples. Furthermore, ages determined by crater-counting techniques have indicated the presence of lavas perhaps as young as 2 billion years. Thus, although now long absent, the majority of volcanic flows on the Moon took place over a time span of about 1 billion years. Additionally, crater-counting data indicate that individual maria may have witnessed eruptions for a similar time span, the older materials being deeply buried near the basins' centers.

The volcanic plains of the Moon, called "seas" or (Latin) "maria," can be seen in this image from Apollo 15. These smooth surface areas form nearly one-fifth of the surface. (NASA)

Crater counting and superposition relationships have led to the construction of a relative timescale based on the time of formation of large impact craters and basins. Subdivisions of this timescale in order of decreasing age are pre-Nectarian, Nectarian, Early Imbrian, Late Imbrian, Eratosthenian, and Copernican. Two-thirds of the nearside surface of the maria is of Late Imbrian age; much of the remainder is of Eratosthenian age. Early Imbrian, Nectarian, and pre-Nectarian lavas could have been buried by impact ejecta and younger lavas; Copernican-aged lavas are restricted to small areas of the western near side.

Morphological evidence for volcanism is found in the form of lava-flow fronts, sinuous rilles, mare domes, cones, and pyroclastic deposits. Measured flow fronts have heights which average about 30 meters but range from 10 to 63 meters. It is highly probable that thinner flows once existed, but they were subject to obliteration by more than 2 billion years of erosion from meteorite impacts. Flow fronts outline individual eruptions more than several hundred kilometers in length. The great size of these features is attributable to the extremely low viscosities of erupted materials and their high-volume output rate. Sinuous rilles superficially resemble terrestrial river channels and range from hundreds of meters to 3 kilometers in width and a few kilometers to 300 kilometers long. These features are believed to be lava channels or collapsed lava tubes. The Apollo 15 mission included exploration at the edge of one of these structures, Hadley Rille. More than three hundred mare domes have been identified on the Moon. They have shapes and dimensions comparable to small terrestrial shield volcanoes.

Conical structures are common in the Marius Hills volcanic complex within Oceanus Procellarum but are relatively rare elsewhere. It is believed that these structures are the lunar equivalent of a terrestrial Strombolian eruption style. The Marius Hills area had been seriously considered as a landing site for Apollo astronauts, but it did not make the cut when three Apollo missions were canceled as a result of budget cuts.

Pyroclastic deposits on the Moon can be subdivided into two major groups: dark halo deposits and regional dark mantle deposits. Dark halo deposits extend to ranges of about 5 kilometers from an endogenic crater. Dark mantle deposits cover areas of up to 40,000 square kilometers. The Apollo 17 mission included the only geologist astronaut, Harrison "Jack" Schmitt, to fly to the Moon to do fieldwork on its surface. On that final lunar landing mission, the astronauts returned samples of this pyroclastic material in the form of orange "soil" found in the Taurus-Littrow valley. Stratigraphy and compositional variations of lunar pyroclastic materials suggest an age range comparable to the mare basalts.

Compositionally and texturally, lunar mare samples returned to Earth are basaltic lavas and glasses. Samples from each landing site are unique in terms of their major and minor element chemistry. Additionally, different chemistries can be recognized in the basalts at each site. General distinctions have been drawn between high, intermediate, low, and very low titanium basalts and feldspathic basalts. It must be remembered that samples have been returned from only six mare sites. Fortunately, basalts are rich in the transition group metals, particularly iron and titanium. It has been possible to map the distribution of compositionally different basaltic materials across the nearside of the Moon on the basis of their spectral characteristics. More than one dozen spectrally different units have been identified. There appear to be no simple age-composition relationships or composition-location relationships. For example, there are both old and young titanium-rich basalts, and titanium-rich basalts in both the eastern and the western hemispheres of the Moon.

There are also dome-shaped features on the lunar surface whose shapes and spectral characteristics serve to distinguish them from mare domes. These features have heights of up to slightly more than 1 kilometer and areas of up to 500 square kilometers. They are all close to the highland-mare boundary in Oceanus Procellarum. Their spectra have suggested a comparison with KREEP basalts, which are rich in potassium, rare Earth elements, and phosphorus. If of volcanic origin, their shapes would indicate formation by higher-viscosity materials or at lower volumetric eruption rates than mare basalts. Stratigraphic analysis indicates that they were emplaced at the lunar surface at the same time that basaltic eruptions were taking place elsewhere.

In summary, volcanic activity, perhaps dominated by KREEP-rich materials, preceded development of younger large impact basins. Cavities created by the large impacts became the sites of basalt deposition, and deeper, central portions of the basins were the first to be filled. Widespread, flood-type volcanism was locally accompanied by the eruption of more volatile-rich pyroclastic material. With progressive infilling, the magma had an increasingly difficult task in penetrating the thick, high-density basin fill, so volcanism shifted to the periphery of the basins, found outlets in the floors of circumbasinal craters, and overflowed into the surrounding terrains.

METHODS OF STUDY

Locations and stratigraphic relations of the lunar maria have been determined by geologic mapping using Earth-based and orbital photographs as the database. Relative ages of different surface units have been determined by measuring the number of impact craters of a given size within a specified area or by determining the extent to which craters of a given size have been modified with respect to fresh, similarly sized craters. Both of these techniques rely on the fact that the Moon is under constant bombardment by meteorite particles, so that the longer an area has been exposed at the lunar surface, the more craters it will have and the more the pre-impact surface will be affected by erosion and deposition. In many cases it is possible to identify secondary craters created by material excavated by a primary impact. These secondary craters (and the primary crater) have to be younger than the material into which they impacted. By mapping these relationships, scientists have established the lunar stratigraphic column.

The advantage of the return of lunar samples was that the relative ages calculated by the cratering techniques could be related to the absolute ages determined by radiometric dating. Furthermore, exposure times (rather than crystallization or metamorphic ages) of the returned samples at the lunar surface have been established by determining their amounts of solar-derived particles. Sizes of such structures as lava-flow fronts have been determined by measuring the lengths of shadows and through the use of laser altimeter carried aboard orbiting spacecraft. These data have been compiled to make topographic maps.

Many techniques have been employed to determine the thickness of the mare basalts. One geophysical technique relies on the fact that the mare basalts have a greater density than the surrounding highlands materials. This greater density results in the maria exerting a greater gravitational pull on a spacecraft, causing minute changes in its orbital motions. This phenomenon led to the discovery of strong pulls over the younger circular impact basins and the formulation of mascons (short for mass concentrations). Morphometric techniques for determining mare thicknesses rely on the fact that impact craters have regular and predictable dimensional characteristics. For example, a fresh, bowl-shaped crater has a depth equal to one-sixth its diameter. If an impact crater has penetrated a mare surface and excavated pre-mare materials from beneath, then the mare must be less thick than one-sixth of the crater diameter. Similarly, if all pre-mare craters with less than a specific diameter have been buried by lava, then, again, a minimum depth can be established. At the larger end of the scale, it is possible to take a topographic map of a relatively unflooded young basin, such as the Orientale Basin, and artificially raise the lava level parallel to the contours until the basin looks like the more deeply flooded basins. Height difference is then a measure of the mare infill thickness in the more deeply flooded basin.

Composition of mare basalts has been established by standard geochemical techniques applied to returned samples. More Moon-wide data have been obtained from orbital geochemi-

Lunar Maria

Name	*Latitude*	*Longitude*	*Diameter*	*English Name*
Mare Anguis	22.6N	67.7E	150.0	Serpent Sea
Mare Australe	38.9S	93.0E	603.0	Southern Sea
Mare Cognitum	10.0S	23.1W	376.0	Sea That Has Become Known
Mare Crisium	17.0N	59.1E	418.0	Sea of Crises
Mare Fecunditatis	7.8S	51.3E	909.0	Sea of Fecundity
Mare Frigoris	56.0N	1.4E	1,596.0	Sea of Cold
Mare Humboldtianum	56.8N	81.5E	273.0	(Alexander von Humboldt, German natural historian, 1769-1859)
Mare Humorum	24.4S	38.6W	389.0	Sea of Moisture
Mare Imbrium	32.8N	15.6W	1,123.0	Sea of Showers
Mare Ingenii	33.7S	163.5E	318.0	Sea of Cleverness
Mare Insularum	7.5N	30.9W	513.0	Sea of Islands
Mare Marginis	13.3N	86.1E	420.0	Sea of the Edge
Mare Moscoviense	27.3N	147.9E	277.0	Sea of Muscovy
Mare Nectaris	15.2S	35.5E	333.0	Sea of Nectar
Mare Nubium	21.3S	16.6W	715.0	Sea of Clouds
Mare Orientale	19.4S	92.8W	327.0	Eastern Sea
Mare Serenitatis	28.0N	17.5E	707.0	Sea of Serenity
Mare Smythii	1.3N	87.5E	373.0	(William Henry Smyth, British astronomer, 1788-1865)
Mare Spumans	1.1N	65.1E	139.0	Foaming Sea
Mare Tranquillitatis	8.5N	31.4E	873.0	Sea of Tranquillity
Mare Undarum	6.8N	68.4E	243.0	Sea of Waves
Mare Vaporum	13.3N	3.6E	245.0	Sea of Vapors

Source: National Space Science Data Center, NASA Goddard Space Flight Center.

cal experiments. Information on radon and polonium variations was gathered by alpha-particle spectrometry; uranium, potassium, and thorium concentrations, and the elemental abundances of oxygen, silicon, iron, magnesium, and titanium were determined by gamma-ray spectrometry; and aluminum, silicon, and magnesium variations were determined using X-ray fluorescence data. Various photographic and reflectance spectroscopy techniques have been employed by Earth-based observers to determine transition element variations, mineral compositions, and glass contents of the mare surfaces facing the Earth.

Lunar studies ceased after the Soviet Luna 24 mission robotically returned to Earth samples from the Sea of Crises (Mare Crisium) in 1976. The Moon was ignored for nearly two decades in terms of planetary science programs involving robotic spacecraft. Then in the 1990's the Ballistic Defense Organization sent the Clementine spacecraft to test new sensors and in the process map lunar resources. NASA followed with the more capable Lunar Prospector. Clementine and Lunar Prospector returned tantalizing suggestions of water ice on the surface of the Moon in permanently shadowed areas. Data from the Clementine spacecraft in 1994 appear to indicate an ice field near the Moon's south pole. Five to ten meters deep and sixteen thousand square kilometers in area, the ice field is mixed with soil; it could be a valuable resource for crewed lunar bases. That finding was substantiated by NASA's Lunar Prospector spacecraft, which orbited the Moon from January, 1998, to June, 1999, and determined the chemical composition of the lunar surface using alpha particle, neutron, and gamma-ray spectrometers. Renewed interest in the Moon followed. The Chinese, Indian, European, and Japanese space programs sent spacecraft into orbit about the Moon in the first decade of the twenty-first century, and NASA prepared to launch the Lunar Reconnaissance Orbiter in 2009.

In the aftermath of the *Columbia* accident (2003), NASA was directed by the Bush administration to fulfill its Space Vision for Exploration, which calls for a return to the Moon during which astronauts would begin setting up a lunar base. Although the plan, the Constellation program, calls for the base to be established near the Moon's south pole, the infrastructure is to be designed so that exploration of the surface far from the base is possible; this could include sorties to the lunar maria. In any event, with China planning to send its taikonauts (the Chinese equivalent of astronauts) to the Moon as well, activity in orbit and on the surface of the Moon should increase dramatically, increasing our understanding of the Moon's history and geology.

CONTEXT

Maria form the least rugged terrain on the Moon, so they are conducive to the safe landing of spacecraft. These flat areas can also facilitate lunar exploration through the use of surface craft. They are most abundant on the Earth-facing hemisphere, where continuous telecommunications are possible with the Earth. Excavation of the maria to produce dwellings would not be very difficult. Alternatively, natural shields to solar radiation, such as lava tubes, could be turned into habitation sites. Several techniques have been proposed for the mining of mare basalts to be used as raw materials for the construction of spacecraft in Earth orbit and to provide sustenance for lunar inhabitants. These factors make the maria primary candidates for the establishment of a permanent, crewed lunar base.

A farside mare site would be an ideal location for the construction of an astronomical observatory. Communications with Earth-based stations would necessitate a more complex satellite system than that required by a nearside base, but at the same time, the shielding from terrestrial electromagnetic radiation provided by the bulk of the Moon would enable clear views of distant galaxies.

From a scientific viewpoint, the surface materials of the maria provide a diary of solar activity and small meteorite impacts extending over the past few billion years. On Earth, by contrast, this information is less available because it has been removed from rocks by erosion, deposition, and the recycling of oceanic plates. Detailed study of the maria can therefore furnish information about what was taking place in the near-Earth solar system from the time of emplacement of the oldest surviving terrestrial rocks through the development of unicellular organisms to the arrival of humans. These data cannot be obtained from the surface of Venus, because of the thick atmosphere of that planet; nor will Mars serve, because of the effects of wind erosion, ice formation, and, perhaps, past erosion by river systems. Therefore, the Moon is a unique natural laboratory, and the maria provide favorable sites for a laboratory inhabited by humans. With the declaration of the Space Vision for Exploration, some of these grand plans may come to fruition during the twenty-first century.

James L. Whitford-Stark

FURTHER READING

Baldwin, Ralph B. *The Measure of the Moon*. Chicago: University of Chicago Press, 1963. A classic book used by those directly involved with the scientific planning of the crewed and uncrewed lunar landings. Surprisingly much of this tour-de-force study of the Moon remains valid. Most valuable for the correlations it draws between impact-cratering phenomena and human-made explosions. Suitable for advanced high school and college students.

Beattie, Donald A. *Taking Science to the Moon: Lunar Experiments and the Apollo Program*. Baltimore: Johns Hopkins University Press, 2003. Explains the science gleaned from the Apollo lunar landings, including the Apollo Lunar Surface Science Experiment Packages (ALSEPs) and their results.

Lunar and Planetary Institute, Houston, Texas. *Basaltic Volcanism on the Terrestrial Planets*. Elmsford, N.Y.: Pergamon Press, 1981. An indispensable text for any study of planetary geology. Entire chapters are devoted to petrology and geochemistry, remote sensing, surface morphologies, radiometric dating, and crater studies. For college-level readers.

Mutch, Thomas A. *Geology of the Moon*. Rev. ed. Princeton, N.J.: Princeton University Press, 1972. A readable textbook for advanced undergraduate and graduate students. Describes the Moon from a stratigraphic viewpoint, with the Apollo results forming the last chapter.

Schmitt, Harrison J. *Return to the Moon: Exploration, Enterprise, and Energy in the Human Settlement of Space*. New York: Copernicus Books, 2006. A scientific and economic plan for prolonged exploration and exploitation of lunar resources, written by the only geologist to land on the Moon and perform field geology and collect documented samples for return to Earth. Discusses the use of helium-3 mining to supply future energy production systems on Earth and thereby finance and systematically expand lunar settlements and research posts.

Schrunk, David, et al. *The Moon: Resources, Future Development, and Settlement*. New York: Springer Praxis, 2007. Examines what is needed to develop human settlements on the lunar surface, beginning with renewed robotic exploration. Explains the benefits of a robust spacefaring civilization.

Schultz, Peter H. *Moon Morphology: Interpretations Based on Lunar Orbiter Photography*. Austin: University of Texas Press, 1974. A post-Apollo photographic encyclopedia of lunar surface features. Aimed primarily at a research-level audience.

Spudis, Paul D. *The Geology of Multi-ring Impact Basins: The Moon and Other Planets*. New York: Cambridge University Press, 1993. A comprehensive geological study of large impact craters formed on the planets and their satellites. Well illustrated with photographs and diagrams. Suitable for both the lay reader and scientific audiences.

Tumlinson, Rick N., and Erin Medlicott, eds. *Return to the Moon*. New York: Collector's Guide Publishing, 2005. A series of essays by lunar experts that examine the Space Vision for Exploration. A broad examination of a return to the Moon, this time to stay and set up lasting infrastructure.

Wilhelms, Don E. *The Geologic History of the Moon*. Professional Paper 1348. Denver, Colo.: U.S. Geological Survey, Books and Open-File Reports Section, 1987. This volume synthesizes the author's vast accumulated knowledge of mapping the lunar surface with information derived from every other field of lunar research. An essential source for budding lunar scientists and a reference source for specialists needing to put their work in perspective. Heavily illustrated.

_______. *To a Rocky Moon: A Geologist's History of Lunar Exploration*. Phoenix: University of Arizona Press, 1994. Numerous books portray the Apollo astronauts, who successfully completed six lunar landings, with heroic prose. Others describe the science learned from Apollo investigations and returned Moon rocks and soil samples. This book is a unique geology-driven explanation for the way individual missions were planned and carried out for maximum science return. Includes descriptions of real-time evaluations of field geology aspects on each Apollo mission.

See also: Earth-Moon Relations; Impact Cratering; Lunar Craters; Lunar History; Lunar Interior; Lunar Regolith Samples; Lunar Rocks; Lunar Surface Experiments; Meteoroids from the Moon and Mars; Planetary Satellites; Solar System: Origins.

Lunar Regolith Samples

Category: Natural Planetary Satellites

Initial study of lunar soils focused on ensuring the safety of the crewed Apollo spacecraft upon landing. Returned soil samples have since been analyzed to determine the origin and evolution of the Moon. Future studies will assess the suitability of lunar resources for utilization in construction of a Moon base and as raw materials for space manufacturing.

OVERVIEW

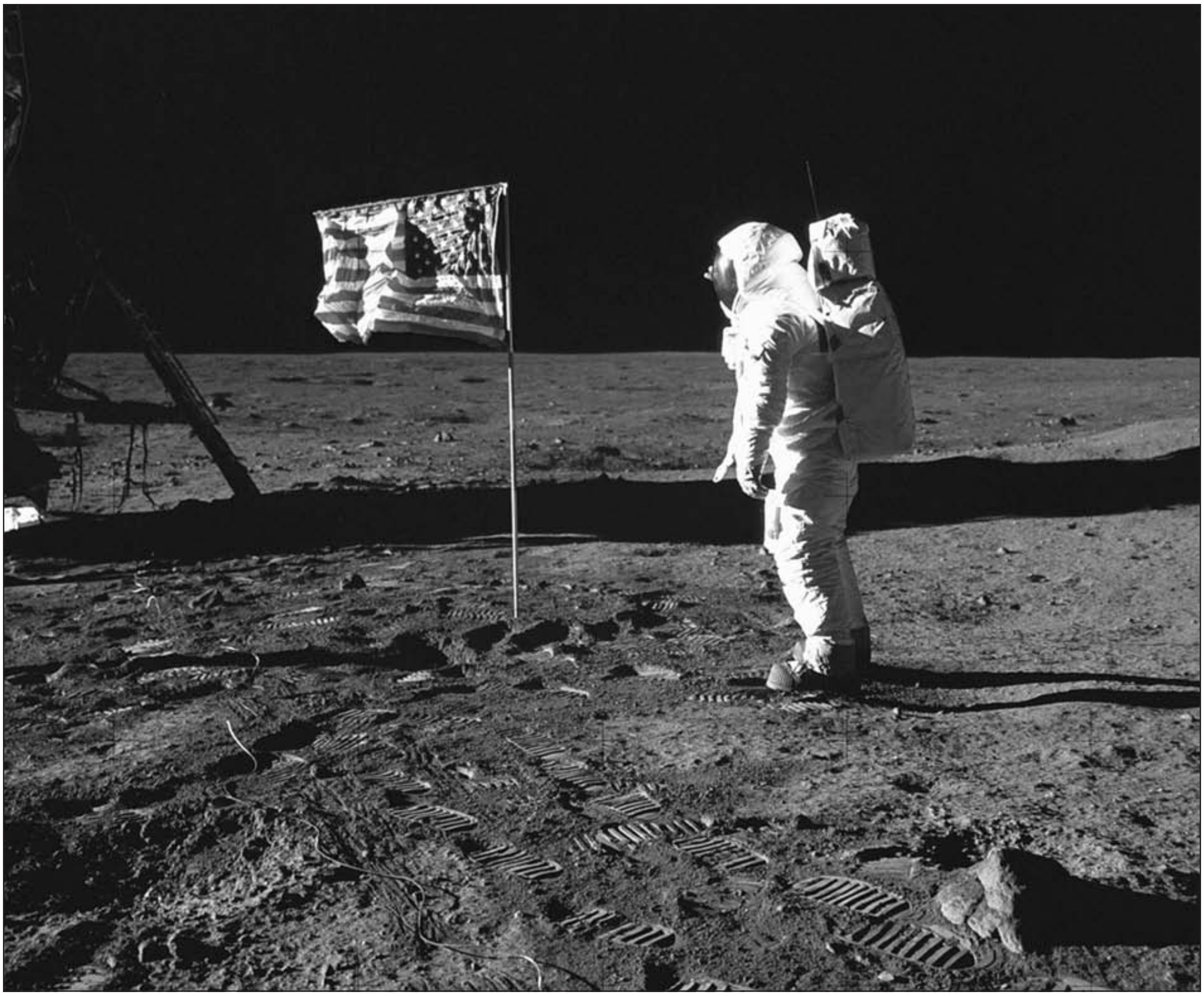
Existence of a "soil," or layer of small particles covering the lunar surface, was inferred prior to the first spacecraft's landing on the Moon. At full Moon, the lunar surface is observed to be bright from edge to edge, exhibiting only minimal "limb darkening" (a decrease in intensity of the light reflected near the edges of a smooth sphere). This led early observers to conclude that the uppermost surface layer of the Moon was porous on the centimeter scale, suggesting a surface dust layer. Determination of the thickness and physical properties of this dust layer was important to the success of the crewed lunar landings. Cornell astronomer Thomas Gold had postulated that the Apollo Lunar Module might sink in a thick surface dust layer that could not bear the weight of the vehicle.

Early in situ investigations of the properties of the lunar soils were conducted by lunar soft-landing spacecraft in the Soviet Luna series and the Surveyor spacecraft series launched by the National Aeronautics and Space Administration (NASA). On February 3, 1966, Luna 9, the first spacecraft to soft-land on the Moon, returned panoramic photographs of the surface and demonstrated that the soil was firm enough to support the 100-kilogram payload without noticeable effect. Surveyor spacecraft and the Luna 13, 17, and 21 soft-landers carried instruments that determined the lunar soil's composition and physical properties.

Nevertheless, the return of lunar samples to Earth for laboratory analysis offered significant advantages; earthbound instruments were not limited, as were the lightweight ones suitable for space flight. The first samples of lunar soil were returned to Earth by Apollo 11. During their stay on the lunar surface, the Apollo 11 astronauts collected 22 kilograms of lunar material. Of this sample, 11 kilograms were categorized as "fines" (particles smaller than 1 centimeter).

In addition to surface soil samples, Apollo astronauts obtained "core samples," cylindrical samples of lunar soil taken by pushing a tube vertically into the lunar surface. Preserved layers in cores provide information on the rates of depositing and mixing of the soil. Apollo 11 astronaut Edwin E. Aldrin collected two core samples in 2-centimeter-diameter tubes pushed down into the lunar surface. The first core, about 10 centimeters long, contained 51 grams of material, while the second, measuring 13.5 centimeters, weighed 65 grams.

Apollo 11 lunar samples were returned to Earth on July 24, 1969, and flown to the Manned Spacecraft Center (MSC, later re-



Apollo 11 delivered the first human beings to the surface of the Moon in 1969 and initiated a series of missions, several of which returned regolith samples. Here astronaut Buzz Aldrin salutes an American flag, left on the surface. The flag had to be backed by a structure that would make it appear to be waving in a nonexistent lunar wind. (NASA)

named Johnson Space Center), in Houston, Texas. There the samples were placed in quarantine in the Lunar Receiving Laboratory (LRL) for a period of one month while biological analyses were conducted. The Lunar Sample Preliminary Analysis Team (LSPAT), consisting of MSC scientists and visiting scientists, was permitted to study samples under controlled conditions during the quarantine period. They first exposed small chips from the samples to nitrogen, oxygen, and air at various humidities to ensure that laboratory analysis conditions did not cause adverse reactions or sample deterioration. Within the LRL, samples were characterized by mineralogical and chemical techniques. In addition, experiments were performed to determine the effects of cosmic-ray exposure on the lunar material, the organic carbon content, and the noble gas concentrations. Additional experiments searched for "magnetic monopoles" (particles of isolated magnetic charge whose existence is postulated by elementary particle physicists).

Following quarantine, lunar samples were made available to about 110 scientists, selected by the Office of Space Science and Applications, to perform a variety of experiments. These scientists—from twenty-one universities, two industrial facilities, three private institutions, and ten government laboratories—included twenty-seven scientists from England, Germany, Canada, Japan, Finland, and Switzerland. Because of the fineness of the returned material, the scientists developed new techniques and instruments to perform a variety of experiments on the lunar samples. After six months, they met to discuss their results at the Apollo 11 Lunar Science Conference, held at MSC from January 5 to January 8, 1970. The Apollo 11 Lunar Science Conference evolved into an annual meeting at which scientists from around the world report their latest results on lunar sample research and planetary science. NASA continues to allocate samples collected during the Apollo missions as new instruments or new techniques warrant further experiments.

Fines were shown to consist of a mixture of glassy materials and small crystal fragments. Within the cores, the majority of the particles ranged in size from 1 millimeter down to 30 micrometers. Glasses exhibited a variety of colors from pale or colorless to gray to wine red, orange, green, brown, and yellow. Crystal fragments were dominated by the minerals plagioclase, clinopyroxene, ilmenite, and olivine.

Lunar material from five additional sites on the Moon was returned to Earth by the Apollo 12, 14, 15, 16, and 17 missions. Core samples up to 2.6 meters in length were obtained. More than 380 kilograms of lunar material were returned to Earth by the Apollo program, more than half of which was collected on the final two Apollo flights.

The Soviet Union's recovery of lunar materials involved the uncrewed spacecraft in the Luna series. Luna 16, which landed on the Moon on September 20, 1970, returned a single 35-centimeter drill core containing 101 grams of soil on September 24. Luna 20 returned a similar sample in February of 1972. The more advanced Luna 24 spacecraft, which landed in the Sea of Crises area of the Moon on August 18, 1976, returned a two-meter core sample.

These lunar samples have been made available to scientific investigators throughout the world. An agreement between the National Aeronautics and Space Administration (NASA) and the Soviet Academy of Sciences in 1971 provided for the exchange of Apollo and Luna samples, allowing all investigators to have access to the soils collected at the nine lunar sites sampled by either Apollo or Luna missions.

Proposals to establish a lunar base and to develop space manufacturing facilities have focused attention on the lunar soils as construction material and as raw material for the manufacturing process. Procedures for the extraction of aluminum from the plagioclase, titanium from the ilmenite, and magnesium from the olivine grains in the soils have been described. As a by-product of these extractions, silicon and oxygen can also be produced. Although lunar soils contain essentially no water, hydrogen implanted by the solar wind might be obtained from the ilmenite and then be used to react with oxygen to produce water. Use of implanted hydrogen as a rocket fuel has also been proposed. Thus, it appears that the high cost of launching materials from Earth's surface for space manufacturing can be circumvented by acquiring the bulk of the raw materials in the low-gravity environment of the Moon.

KNOWLEDGE GAINED

Lunar regolith (a blanket of broken fragments ranging from dust- to meter-sized blocks of rock) was found to cover almost all the Moon to an estimated depth of a few meters. The fine component of the regolith, generally referred to as soil, was shown to be very different from terrestrial soil. On Earth, the soil is formed by the complex action of atmospheric weathering and biological activity. Lunar soil is produced by bombardment of the surface rocks by meteorites and micrometeorites, producing craters and microcraters; the impact debris is then scattered over the lunar surface.

Chemical composition of the soil was discovered to differ slightly from that of the lunar rocks. By subtraction of the rock composition from that of the soil, the chemical composition of the added component was found to be similar to the composition of primitive stone meteorites that fall on Earth. Thus, meteoritic bombardment was proved to be an important alteration process on the Moon.

Analyses of the layered structure of the tube samples suggested that large meteorite impacts have thrown layers of soil across the lunar landscape. The top surface of this layer has been stirred by impacts of more numerous micrometeorites. The layer structure observed in long core tubes collected on later Apollo missions agrees well with computer simulations of the rate of soil mixing, calculated assuming the rate at which small particles strike the Moon has been approximately equal to the measured current rate for millions of years. Radiation damage in grains of the lunar soil, as well as in lunar rocks, suggests that the rate of emission of heavy ions by the Sun has also been relatively constant over the past few million years.

Layers in the lunar soil record the history of bombardment of the Moon, with a single core providing samples of material ejected from craters far apart on the lunar surface. One Apollo 12 core contains a layer, light gray in color and rich in silica, that has been identified as ejecta (material thrown out by volcanic eruptions) from the crater Copernicus, 75 kilometers from the core site. Thus, Apollo core samples have provided information on the composition and mineralogy of locations far from the sampling site.

CONTEXT

Both the Apollo and Luna programs returned samples of the lunar suface material. Though these lunar rocks could be compared to similar terrestrial rocks, lunar soil was clearly different from its terrestrial counterpart. Analysis of soils and the stratigraphy of core samples provided information necessary to confirm the importance of meteoritic bombardment in producing rock fragments which make up lunar soil and in stirring the upper layer of the soil once it is placed on the lunar surface. Measurement of effects of cosmic rays in an Apollo 15 core sample indicated that layered structures within the lunar soil can be preserved undisturbed on the lunar surface for periods of 500 million years. Cores established that the lunar erosion process proceeds at a rate of between 1 and 2 millimeters per million years, about one thousand times slower than on Earth.

The depth of the lunar soil layer confirmed that the rate at which meteorites and micrometeorites are hitting the Moon has remained essentially unchanged over the past four billion years. The presence of ion damage caused by solar flares in soils buried for millions of years also confirmed that the flux of charge particles from the Sun has been relatively constant.

The chemical composition of the bombarding material was also deduced from the difference between soil and bulk rock compositions. Several distinct compositions were observed, but all were generally similar to the known meteorites collected on Earth.

Studies of returned lunar soils also demonstrated their potential as a raw material for lunar base construction and for space manufacturing. Proposals for lunar mining have generally assumed the soils as the starting material. Since the delivery of large masses of material from the lunar surface to near-Earth orbit requires less energy than delivery of the same mass of material from the surface of Earth, lunar soils are expected to play a significant role in the industrialization of space. The Apollo and Luna soil samples are available on Earth to test and perfect extraction techniques in Earth laboratories so that equipment appropriate for the first lunar mining facility can be designed.

No soil samples were returned to Earth after the Soviet Luna 24 mission in 1976, but in the 1990's a pair of spacecraft, Clementine and Lunar Prospector, mapped the surface for mineral content. A renewed interest in the Moon led the European Space Agency (ESA), Japanese Aerospace Exploration Agency (JAXA), the Chinese National Space Administration (CNSA), and the Indian Space Research Organization (ISRO) to send probes to the Moon in the first decade of the twenty-first century. These spacecraft were named SMART 1 (ESA), Kaguya (JAXA), Chang'e 1 (CNSA), and Chandrayaan 1 (ISRO).

In 2009, NASA was preparing to launch the Lunar Reconnaissance Orbiter to begin a comprehensive mapping of lunar resources and assist in the determination of the best location near the Moon's south pole for the construction of a permanently occupied lunar base. That goal was the first objective of the Bush administration's Vision for Space Exploration set up in the aftermath of the 2003 *Columbia* accident.

NASA was charged with a return of astronauts to the Moon by 2019, the fiftieth anniversary of the Apollo 11 lunar landing. However, the Chinese also stated a firm commitment to evolve their early Shenzhou piloted missions into eventual flights to the Moon, and there is a very real chance that when American astronauts return to the Moon, Chinese taikonauts (the Chinese equivalent of astronauts) will already have visited the lunar surface. Regardless of which nations return to the Moon, the renewed interest in lunar exploration portends advances in understanding of our solar system and a first step in lunar colonization.

George J. Flynn

FURTHER READING

Baldwin, R. A. *The Measure of the Moon*. Chicago: University of Chicago Press, 1963. The first nine chapters outline the characteristics of both lunar and terrestrial, human-made, and natural craters. Written in pre-metric-system times, this classic text requires mathematical (though simple) conversions. Suitable for advanced high school and college students.

Beattie, Donald A. *Taking Science to the Moon: Lunar Experiments and the Apollo Program*. Baltimore: Johns Hopkins University Press, 2003. Explains the science gleaned from the Apollo lunar landings, including the Apollo Lunar Surface Science Experiment Packages (ALSEPs) and their results.

French, Bevan M. *The Moon Book*. New York: Penguin Books, 1977. A classic work that describes the origin of the Earth-Moon system and evaluates these explanations in the light of the lunar-sample results from the Apollo and Luna missions.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Particularly strong in its presentation of Earth-Moon science. Uses the comparative planet approach, which means there are sections on each of the planets in chapters such as "Planetary Interiors," "Planetary Surfaces," and "Planetary Atmospheres."

Jefferis, David. *Return to the Moon*. New York: Crabtree Children's Books, 2007. An explanation of contemporary plans to return American astronauts to the Moon as part of Constellation program operations with the Orion Crew Exploration Vehicle and Altair landing craft. Aimed at a younger audience.

King, Elbert A., Jr. *Space Geology: An Introduction*. New York: John Wiley and Sons, 1976. This 349-page text, intended for advanced students and written by the curator of NASA's Apollo 11 Lunar Receiving Laboratory, includes a well-illustrated section describing the analyses of lunar soils.

Levinson, Alfred Abraham, ed. *Apollo 11 Lunar Science Conference: Proceedings*. 3 vols. Elmsford, N.Y.: Pergamon Press, 1970. This 2,490-page work includes articles by the Apollo 11 principal investigators that report the techniques employed and the preliminary results obtained in their analyses of the Apollo 11 lunar samples.

Mason, Brian, and William G. Melson. *The Lunar Rocks*. New York: Wiley-Interscience, 1970. Written shortly after the Apollo 12 sample return, this book analyzes the results reported on the Apollo 11 and 12 samples. Includes a chapter describing the fines (lunar soils).

Schmitt, Harrison J. *Return to the Moon: Exploration, Enterprise, and Energy in the Human Settlement of Space*. New York: Copernicus Books, 2006. A scientific and economic plan for prolonged exploration and exploitation of lunar resources written by the only geologist to land on the Moon. Talks about helium-3 mining, aimed at supplying future energy production systems on Earth and thereby financing and systematically expanding lunar settlements and research posts.

Taylor, S. R. *Planetary Science: A Lunar Perspective*. Houston: Lunar and Planetary Institute, 1982. This well-illustrated college-level work, written by a member of the Lunar Sample Preliminary Analysis Team for Apollo 11, describes the geological processes active on moons and planets in the light of the discoveries from the Apollo lunar samples.

Wilhelms, Don E. *To a Rocky Moon: A Geologist's History of Lunar Exploration*. Phoenix: University of Arizona Press, 1994. Numerous books portray the Apollo astronauts, who successfully completed six lunar landings, with heroic prose. Others describe the science learned from Apollo investigations and returned Moon rocks and soil samples. This book is a rather unique personal account in that it is a geology-driven explanation for the way individual missions were planned and carried out for maximum science return. Includes descriptions of real-time evaluations of field geology aspects on each Apollo mission.

Wood, John A. "The Lunar Soil." *Scientific American* 223 (August, 1970): 14-23. In this article, one of the principal investigators of the Apollo lunar samples describes the results of the first analyses of the returned lunar soil samples. Well illustrated and intended for the general audience.

See also: Earth-Moon Relations; Impact Cratering; Lunar Craters; Lunar History; Lunar Interior; Lunar Maria; Lunar Rocks; Lunar Surface Experiments; Meteoroids from the Moon and Mars; Planetary Satellites.

Lunar Rocks

Category: Natural Planetary Satellites

Lunar rocks are among the materials brought to Earth from the surface of the Moon during the Soviet robotic Luna and the American Apollo crewed Moon missions of the 1960's and 1970's. These samples were analyzed by international teams of scientists in an effort to enhance understanding of the physical nature of the Moon and its origins. All samples were cataloged according to criteria relating to mineralogy, crystallography, geochronology, geochemistry, magnetism, radioactivity, and other characteristics.

OVERVIEW

When direct exploration of the surface of the Moon began during the 1960's, scientists around the world looked forward with great anticipation to the study and analysis of materials brought back to Earth from the lunar surface by the crewed space missions of the Apollo program and the uncrewed programs of the Soviet Union such as the Luna missions. As new and different material arrived at the completion of each successive mission, an astonishing wealth of information about the origin and physical character of the Moon began to accumulate. A great deal of the Apollo lunar samples remain to be studied even after forty years or more. A conscious effort had been made to save large portions of the Moon rock collection for new technologies that could be employed to study them in ways not possible back in 1969 through 1972, the era of the Apollo lunar landings.

Using the latest scientific techniques to analyze soil and rock gathered during the missions, researchers were able to identify the basic elements and minerals that constitute the Moon's surface. Using that knowledge in combination with other information gathered by remote-sensing technology and sophisticated photography techniques over several decades, they were able to extrapolate information about the origin of the Moon. Perhaps most important, they were able to determine with some accuracy a chronology of events in lunar evolution. While there remain many unsolved mysteries about the Moon and its history, it is remarkable how much information was gathered from the study of so-called Moon rocks.

"Moon rocks" are defined as surface materials that exceed 1 centimeter in diameter. Anything smaller is considered lunar "soil," or regolith, although lunar regolith is unlike Earth soil, which usually contains large amounts of decayed organic material and moisture. On the Moon, the lack of atmosphere and organic material gives the soil a composition similar to dry, clean sand. It is probably powdered rock, created when large objects such as meteoroids impacted violently with the surface of the Moon over many millennia.

Rocks on the surface of the Moon are rich in calcium and aluminum. Below 20 kilometers, the surface is not broken as extensively as it is near the top. The upper mantle is from 200 to 300 kilometers thick and contains high concentrations of magnesium and iron silicate, pyroxene, and olivine. There is reason to believe that the core is iron-rich and produced a global magnetic field during the early days of the Moon's history. Most of the radioactive materials on the Moon are located at or near the surface. Two major activities that have resulted in the character of the lunar surface are volcanism and impact events. Lowland regions exhibit evidence that there has been great volcanic activity on the Moon, and across the lunar landscape there is evidence that impacts have resulted in broad redistribution of rock and soil. Occasionally that activity has resulted in the breaking up of bedrock and basalt breccias that formed prior to later redistribution, recrystallization, or both.

Each of the lunar missions that returned to Earth with samples of Moon rock expanded the body of knowledge about the geological character of the Moon's surface. Each mission landed at a different location on the lunar surface to ensure that as broad a collection of materials as possible would be acquired given the limited number of missions planned and the limited payload capacity available on returning spacecraft. For the purposes of lunar exploration, two primary regions of the Moon were identified. Generally speaking, the lunar highlands are regions that appear light on the surface on the Moon. The maria constitute that portion of the lunar surface that is dark. The highlands are generally composed of remnants of ancient volcanic cones and the ridges of giant impact craters. The maria are surface areas characterized by volcanic chamber depressions and impact

Astronaut Harrison Schmitt stands at the foot of an outcrop of boulders near the Taurus-Littrow landing site of Apollo 17. (NASA/E. Cernan) crater basins. Analysis of rocks collected at various spots within each of the two regions demonstrates that each has unique physical characteristics.

Moon rocks are of two main types, crystallized (or igneous) and breccias. The igneous rock encountered on the lunar surface is most often volcanic in origin and appears to have been scattered over wide areas during violent eruptions. Some specimens are believed to have been thrown between 100 and 1,000 kilometers, either during eruptions or as a result of meteor impacts. Certain highland samples contain evidence of the existence of plutonic rock, formed when volcanic material cools below the surface.

The Apollo 17 crew, working with the benefit of knowledge gained in previous crewed and uncrewed missions, further subcategorized lunar rock as follows: basalts; dark matrix breccias; glass-bonded agglutinates; vesicular green-gray breccias; blue-gray breccias; layered, foliated, light-gray breccias; and brecciated gabbroic rocks. The distribution of Moon-rock material is different for highland and mare regions. In the highlands, the surface consists of approximately 10 percent plutonic rock, 85 percent breccias, and 5 percent volcanic material, primarily basalts. Mare regions contain 90 percent volcanic basalts and 10 percent breccias.

It is important to note that the number of minerals identified on the Moon is dwarfed by the number found on Earth; this difference probably results from the Moon's lack of atmosphere, particularly oxygen and moisture, which often interact with elements in the weathering process to form minerals. There are some two thousand minerals on Earth, whereas there are only two hundred known to exist on the Moon. The number of primary minerals, however, is similar on both bodies. The principal minerals discovered on the Moon at the regolith level include clinopyroxene, plagioclase, olivine, ilmenite (by far the most abundant, 15-20 percent by volume), trydymite, cristobalite, and orthopyroxenes. Several new minerals have been identified, including armalcolite, pyroxferroite, and tranquillityite. It is interesting to note that armalcolite was named for the three astronauts on the Apollo mission that returned with it: Neil Armstrong, "Buzz" Aldrin, and Michael Collins. Armalcolite has since been found on Earth; it is formed when crystallization occurs in the absence of moisture and oxygen.

Ancient orange glass beads, thought to have been spewed out during volcanic eruptions, have given scientists a glimpse of what constitutes the core of the Moon's interior. These beads contain high levels of lead, zinc, tellurium, and sulfur and are thought to have originated as deep as 300 kilometers below the surface of the Moon. Lunar rock samples have been individually but not uniformly rich in such elements as aluminum, magnesium, potassium, phosphorus, titanium, iron, chromium, and zirconium. Some lunar rocks are referred to as glassy agglutinates. These rocks are believed to have been formed when volcanic action or meteoric impact caused fine rock dust to weld together, a process known as impact melting. Their appearance ranges from dark opaque to transparent. Other rocks have different origins but are covered by glassy materials thought to have formed and been distributed in the same manner. While the great majority of collected lunar rocks are igneous, there is some evidence that a metamorphosis has occurred as a result of shock settling of soil, caused by meteor impact and vibrations that occur during the course of violent volcanic activity.

Breccias appear to contain material that has been ejected by impacts or volcanic eruption. There are two types: soil breccia and ejecta blanket breccia. Soil breccia is composed of the same materials that exist as soil in a particular region and probably results from the shock of meteor impact. Soil breccia is also characterized by weak cohesion. Ejecta blanket breccias, by contrast, are thought to have been formed when volcanically ejected materials fused together to form a layer of rock that later underwent various forms of thermal metamorphism. The composition of such rock varies with the depth at which it has been found or at which it is thought to have originated.

The primary source of lunar erosion appears to have been impact cratering, although solar wind and cosmic-ray effects have also altered the character and composition of surface materials. Some lunar samples show cosmic-ray bombardment over long periods of time that has resulted in the alteration of the surface of the sample. The same activity has been useful in helping scientists determine how long a particular sample fragment was in the same location and at what level it has resided within the regolith.

Scientists have also determined with some confidence the age of the Moon and the time of the earliest crystallization of Moon rock formations. Isotopic dating techniques indicate that the earliest crystallization occurred more than 4 billion years ago and that most volcanic activity appears to have spanned some 600 million years.

During the Apollo 16 mission, Commander John W. Young took the Lunar Roving Vehicle for a spin near the Descartes landing site; 90 percent of soil and rock samples brought from this region had high concentrations of aluminum. (NASA/Charles M. Duke, Jr.)

METHODS OF STUDY

Before the first direct contact with the surface of the Moon occurred, much was known about the composition and origin of that surface from sophisticated technology capable of analyzing materials remotely. Those techniques included indirect sensing processes involving the visible, infrared, and microwave bands of the electromagnetic spectrum. The understanding derived from this technology was limited until actual samples of surface materials were brought to Earth by the Moon explorers. Scientists could then compare the data from the remote-sensing experiments with direct observations to determine what it was that they were seeing at the time. Radar, which was used in determining the physical nature of the Moon's surface, was successful in establishing that it was loose, a mixture of sand and broken rock. Infrared studies confirmed that the surface was porous and exhibited low thermal conductivity. Photometric and polarimetric studies further refined the knowledge of the character of the surface by establishing that it contains high concentrations of iron-rich basalts (confirmed by the Apollo missions).

Once the first lunar samples arrived back on Earth, scanning electronic microscopes were used to photograph them, an important step in categorizing individual sample fragments. Samples were also sieved to determine the relative abundance of materials in a particular location. In this way, researchers were able to determine where some materials were more abundant than others on the surface when collected using the same procedure. Some samples were cross-sectioned for better observation of stratification, if it existed. Many were tested using standard chemical analysis techniques to determine the chemical nature of the materials. X-ray fluorescence analyses were also conducted to determine the relative concentrations of specific elements or components. Isotopic and carbon dating techniques were used to determine the relative ages of specific samples and to attempt a chronology of the events that have affected the evolution of the materials under study.

CONTEXT

There is enormous scientific interest in the Moon. Many scientists believe that it may hold the key to a better understanding of the physical nature of the solar system—indeed, of the universe itself. Perhaps more important, it may hold the key to a better understanding of the origin of planet Earth and of its relationship to other objects in the universe. Lunar rocks have given scientists the first window into that unknown sphere, helping them to understand the nature of planetary formation. The Luna and Apollo programs demonstrated for the first time that many accepted ideas about the Moon and its composition were accurate, though many others were not. They also demonstrated that the Earth and its satellite have much in common. Except for the lack of a lunar atmosphere of composition similar to that of Earth, they would probably have evolved in much the same way.

Another significant outcome of the study of lunar rock samples has been the unequivocal knowledge that certain elements, metals, and other valuable materials exist in abundance on the Moon and may be important to future generations on Earth. Perhaps more important, however, is the knowledge gained of how specific elements interact under conditions that rarely exist on Earth: namely, with little oxygen and no water. The Moon has now become a giant laboratory that is helping scientists to understand the nature of chemical and geological evolution and order.

In the aftermath of the 2003 Columbia accident, President George W. Bush committed NASA to a Space Vision for Exploration that envisioned the completion of the International Space Station by the second decade of the twenty-first century. At that time the space shuttle fleet would be retired so that humans could resume exploration beyond low-Earth orbit. This plane, which NASA officially named Project Constellation, would have a new crew exploration vehicle (Orion) and lunar lander (Altair) developed to return astronauts to the Moon by 2020, this time to stay and build a permanent base of operations. Constellation efforts were designed to be evolutionary and eventually lead to an expedition to Mars. However, a Moon base would also serve as an outpost from which lunar geology investigations could be organized at a scope well beyond that possible during the Apollo program.

Michael S. Ameigh

FURTHER READING

Beattie, Donald A. *Taking Science to the Moon: Lunar Experiments and the Apollo Program*. Baltimore: Johns Hopkins University Press, 2003. Explains the science gleaned from the Apollo lunar landings, including the Apollo Lunar Surface Science Experiment Packages (ALSEPs) and their results.

Chaikin, Andrew. *A Man on the Moon: The Voyages of the Apollo Astronauts*. New York: Penguin, 2007. A reissue of one of the most engaging accounts of the Apollo program and the exploration of the Moon. Critically acclaimed, for all ages.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through space missions that have investigated all the planets. Takes an astrophysical approach, placing our solar system in a wider context as just one of many similar systems throughout the universe.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Particularly strong in its presentation of Earth-Moon science. Takes a comparative planetology approach rather than including separate chapters for each planet in the solar system.

National Aeronautics and Space Administration. *Preliminary Science Report: Apollo 11*. NASA SP-214. Washington, D.C.: Government Printing Office, 1969. The first humans on the Moon arrived on this mission and conducted preliminary experiments. Well illustrated and full of graphs.

_______. *Preliminary Science Report: Apollo 12*. NASA SP-235. Washington, D.C.: Government Printing Office, 1970. The results of several experiments on this mission advanced our knowledge of lunar geology significantly. Well illustrated and full of graphs.

_______. *Preliminary Science Report: Apollo 14*. NASA SP-272. Washington, D.C.: Government Printing Office, 1971. This Apollo mission returned significant rock samples from the Fra Mauro region. Well illustrated and full of graphs.

_______. *Preliminary Science Report: Apollo 15*. NASA SP-289. Washington, D.C.: Government Printing Office, 1972. This mission examined the Hadley mountain area and was the first to use the Lunar Rover vehicle. The mission's heat-flow experiments suggested a previously molten interior. Well illustrated and full of graphs.

_______. *Preliminary Science Report: Apollo 16*. NASA SP-315. Washington, D.C.: Government Printing Office, 1972. This mission returned more soil samples and again used the Lunar Rover. Well illustrated and full of graphs.

_______. *Preliminary Science Report: Apollo 17*. NASA SP-330. Washington, D.C.: Government Printing Office, 1973. The final Apollo mission to the Moon landed in the Taurus-Littrow Valley and returned the largest collection of lunar rocks to Earth. Well illustrated and full of graphs.

Schmitt, Harrison J. *Return to the Moon: Exploration, Enterprise, and Energy in the Human Settlement of Space*. New York: Copernicus Books, 2006. A scientific and economic plan for prolonged exploration and exploitation of lunar resources, written by the only geologist to land on the Moon, perform field geology, and collect documented samples for return to Earth. Discusses helium-3 mining, which promises to supply future energy production systems on Earth and thereby finance and systematically expand lunar settlements and research posts.

See also: Earth-Moon Relations; Impact Cratering; Lunar Craters; Lunar History; Lunar Interior; Lunar Maria; Lunar Regolith Samples; Lunar Surface Experiments; Meteoroids from the Moon and Mars; Planetary Satellites.

Lunar Surface Experiments

Category: Natural Planetary Satellites

The Moon has been studied from Earth and from lunar orbit, but the most detailed studies must be done on the lunar surface. The United States and the Soviet Union successfully landed eighteen spacecraft on the lunar surface, with many of these missions carrying multiple science payloads.

OVERVIEW

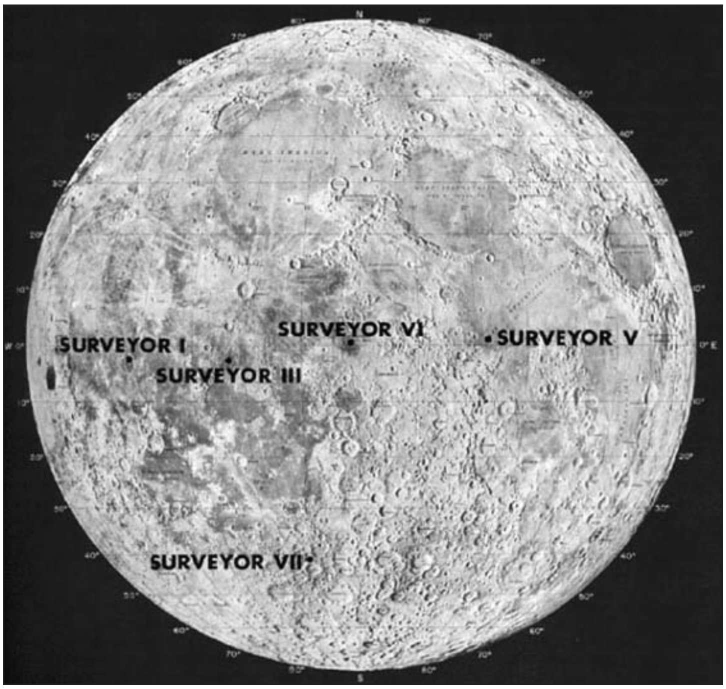
The Soviet Union's Luna 1 spacecraft was the first to visit the vicinity of the Moon, flying past it on January 4, 1959. Later missions by both the United States and the Soviet Union crashed into the Moon, collecting photographic and other data on the way. On February 3, 1966, the Soviets' Luna 9 became the first spacecraft to soft-land on the lunar surface. Luna 9, however, had only cameras and a radiation detector as scientific instruments. Luna 13 landed later that year, carrying instruments to measure the density and strength of the lunar regolith (soil) and to study cosmic-ray reflections from the lunar surface. Also in 1966, on June 2, the United States' Surveyor 1 spacecraft soft-landed on the lunar surface. It, too, carried no more than cameras and landing radar with which to study the Moon, but still returned useful data on the nature of the lunar surface. Four other Surveyor probes landed on the Moon over the next two years. All carried cameras, and two carried robotic arms to scrape the surface in order to determine its consistency and to move the regolith into a better position to photograph it. The final three Surveyors also carried alpha-scattering surface analyzers used to measure the abundances of the many elements making up the lunar regolith.

The Soviet Union successfully landed seven Luna spacecraft on the Moon, with the last one being Luna 24, which touched down August 18, 1976. Three Luna missions (16, 20, and 24) returned small samples of lunar regolith to Earth. Luna 17 and Luna 21 each carried remote-controlled roving vehicles called Lunokhod.

From November, 1970, until September, 1971, Lunokhod 1 traveled 10.5 kilometers and returned nearly twenty thousand images. It carried an X-ray spectrometer to study the composition of the lunar soil. Lunokhod 2 traveled 37 kilometers across the lunar surface from January 16 to June 4, 1973, returning about eighty thousand images and performing mechanical tests of the lunar regolith.

The uncrewed lunar missions have proved very important to our understanding of the Moon, but some of the most important science missions were deployed by astronauts during six Apollo landings from 1969 to 1972. Many of experiments were conducted by several of the missions. In addition to conducting experiments on the lunar surface, the Apollo missions collected nearly 382 kilograms of lunar material to bring back to Earth. These samples included rocks as well as samples of the lunar regolith. The landing sites for the Apollo missions were chosen to provide for samples from a variety of different geological features in order to maximize the impact of only a few sample sites. Scientific instruments carried by Apollo 11 in July, 1969, were solar-powered. However, the Moon's slow rotation means that the lunar night lasts for two weeks. The remaining Apollo missions carried radiothermal generators to supplement power for the instrument packages. Instruments for Apollo 12 through Apollo 17 continued operations until the science stations were shut down on September 30, 1977.

All of the Apollo missions carried seismographic equipment with which to study the Moon below its surface. Most of the seismographs were passive systems, monitoring moonquakes. However, Apollo 14 and Apollo 16 both carried active seismographs. These used mortars to fire explosive shells some distance away from the landing site in order to produce seismic waves that could be used to study the lunar geology. A similar seismic system was also deployed by Apollo 17 in December, 1972, except that it used explosive charges carefully placed by the astronauts rather than mortars. In all three cases, the explosives were not detonated until after the astronauts had left the lunar surface.



Surveyor 1 was the first spacecraft to soft-land on the lunar surface; later Surveyor missions ascertained fundamental surface conditions and paved the way for the Apollo astronauts. (NASA)

Apollo missions also carried lunar dust detectors designed to study the dust disturbed by liftoff of the lunar lander's ascent stage. Later versions of the detector also included provisions to study the long-term degradation of solar panels exposed to the harsh radiation environment of the Moon. Astronauts also set up foils to capture solar wind particles to return to Earth for study. However, in addition to returning particles for study, ion detectors monitored solar wind and cosmic radiation on the lunar surface, and Apollo 12 deployed a spectrometer to measure the composition of the solar wind. These detectors continued to operate for several years after the end of the Apollo program and returned useful data on variations in solar activity. Ion detectors also were used to monitor gas molecules near the lunar surface. These gases constitute what planetary scientists refer to as the lunar atmosphere. Apollo 17 carried a mass spectrometer to measure the composition of the lunar atmosphere.

Supplementing instruments designed to study the Moon, Apollo 16 carried a far ultraviolet telescope and spectrograph. This was the only astronomical instrument placed on the Moon. Among other targets, it was used to study the Earth and the Large Magellanic Cloud.

To study the geophysics of the Moon, the later Apollo missions also carried magnetometers to study the Moon's residual magnetic field, and Apollo 17 carried an experiment to measure the Moon's surface gravity and to monitor it for any variations over time. Heat-flow experiments to measure the amount of heat flowing from the lunar interior were set up on each of the last three Apollo missions. However, astronaut John Young tripped over the cable for the experiment, breaking it. The cable could not be repaired by the astronauts, so the instrument returned no data.

In addition to experiments carried to the Moon to be performed there, several missions carried mirrored corner reflector arrays that were positioned to point back toward Earth. These corner reflectors, completely passive systems, were designed to reflect light striking them back in the direction from which it came. Powerful lasers fired from Earth at the sites of the reflectors are reflected back to Earth, where they can be detected. Careful measurements of the time that it takes for the light to get to the Moon and back are used to determine the lunar distance and variations in the lunar orbit. The astronauts of Apollo 11 deployed the first corner reflector array. Apollo 14 carried another array, and Apollo 15 deployed a much larger array. In addition to the Apollo arrays, both Lunokhod rovers carried such arrays. With the exception of Lunokhod 1's array, which has not reflected lasers since 1971, these corner reflector arrays are still used today in lunar ranging experiments.

KNOWLEDGE GAINED

Though the Moon is our nearest neighbor in space, it is still a long way away from astronomers on Earth, and detailed studies of its surface were not possible until the advent of spacecraft capable of traveling to the Moon. One of the first, and very important, steps in studying the Moon was photographic studies of the surface characteristics. Many of the early surface experiments, particularly those of the Surveyor program, were designed to study the physical properties of the lunar surface to determine if the regolith would be able to support a heavy spacecraft such as the planned crewed missions that were to follow. However, the experimental science stations set up by the Apollo astronauts returned important data about the nature of the Moon itself.

Prior to the surface investigations of the Moon, it was thought that the surface material was largely volcanic in nature. However, the lunar regolith has been found to be composed primarily of impact ejecta from meteorite impacts. Some speculation had been that micrometeorite impacts might have ground the surface of the Moon into a vast ocean of dust, unable to support the weight of a spacecraft landing on it. However that idea was soundly dismissed by the early lunar landings. A layer of fine dust does exist on the lunar surface, but it is more compact than had been thought.

Other surface studies confirmed some ideas put forth after orbital observations that the lunar seas were ancient lava fields that resulted from massive impacts on the lunar surface. Seismological data suggest that the basalts that flooded the impact basins to form the seas did not come all at once, but over a period of time early in the Moon's history. However, a surprising finding was that the impacts that caused these basins occurred primarily near the end of a period of intense bombardment on the Moon rather than randomly distributed in time, as had been suspected.

The experiments performed on the Moon by both crewed and uncrewed missions, along with the studies of lunar samples returned to Earth by the Apollo missions and three Luna missions, have shown the Moon to be an alien world but with some familiarities. The Moon has a much lower density than Earth, and it has a very tiny core. Most of the Moon is composed of material similar to Earth's crust and mantle. Though a few minerals were found on the Moon that did not have counterparts on Earth, most lunar samples were composed of minerals found on Earth. However, the lunar samples were much richer in refractory minerals than Earth rocks, and they contained very few volatile minerals.

Astronomers had assumed that liquid water may have once flowed on the lunar surface, but surface experiments showed a near total lack of water and no indication that water had been present on the Moon. Furthermore, surface investigations suggest that many of the rocks of the lunar highlands formed from a global magma ocean present soon after the Moon's formation.

These findings have revolutionized our understanding of how the Moon and Earth are related to each other. It is now believed, based on the Apollo and Luna findings, that the Moon formed when a giant planetesimal, perhaps the size of the planet Mars, collided with Earth very early in the history of the solar system, perhaps even before the Earth had cooled. The debris from the collision coalesced to form the Moon. Thus, understanding the Moon helps us to understand Earth.

CONTEXT

Much of the early research on the Moon was done in the heat of the space race between the United States and the Soviet Union during the 1960's. However, after initial successes, the political will to continue the study of the Moon faded. The last spacecraft to land on the Moon was Luna 24 in 1976. In the early twenty-first century, however, interest in the Moon revived, and several nations developed plans to land spacecraft on the Moon and once again begin studies on the lunar surface.

Experiments to be done on the Moon in the future will build on the work done in the 1960's and 1970's. Scientific theories about the Moon based on the findings of these earlier experiments have led to several theories of the Moon and its evolution, and upcoming missions will carry experiments to test those theories.

Raymond D. Benge, Jr.

FURTHER READING

Beattie, Donald A. *Taking Science to the Moon: Lunar Experiments and the Apollo Program*. Baltimore: Johns Hopkins University Press, 2001. A well-documented look at the decisions leading to the selection of the lunar science experiments rather than the experiments themselves.

Bond, Peter. *Distant Worlds: Milestones in Planetary Exploration*. New York: Copernicus Books, 2007. This book for the lay reader covers the history of planetary exploration. A chapter is devoted to lunar exploration.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. An excellent college-level introductory astronomy textbook with a chapter devoted to the Moon. The emphasis is on the understanding of the Moon, though, rather than the experiments leading to that understanding.

Hamblin, W. Kenneth, and Eric H. Christiansen. *Exploring the Planets*. New York: Macmillan, 1990. An excellent overview of planetary geology, with an entire chapter devoted to the Moon. It is well illustrated, though all illustrations are in black and white.

Heiken, Grant, and Eric Jones. *On the Moon: The Apollo Journals*. New York: Springer, 2007. An excellent and well-documented resource for studying the Apollo missions, covering each mission, the process behind selecting the landing sites, and the experiments performed.

Mackenzie, Dana. *The Big Splat: Or, How Our Moon Came to Be*. Hoboken, N.J.: John Wiley & Sons, 2003. An excellent explanation of the current, post-Apollo model of how the Moon formed. One chapter covers science done on the Moon. The book has a glossary and a very good bibliography.

Orloff, Richard W., and David M. Harland. *Apollo: The Definitive Sourcebook*. New York: Springer, 2006. A thorough account of the Apollo missions and results, with a glossary and an extensive bibliography.

See also: Earth-Moon Relations; Impact Cratering; Lunar Craters; Lunar History; Lunar Interior; Lunar Maria; Lunar Regolith Samples; Lunar Rocks; Meteoroids from the Moon and Mars; Planetary Satellites.

M

Main Sequence Stars

Category: The Stellar Context

A "main sequence star" is one that is fusing hydrogen into helium in its core. This is the longest-lasting stage in a star's energy-producing life. Most stars, including the Sun, are in this main sequence stage.

OVERVIEW

The Hertzsprung-Russell diagram (or H-R diagram) is a graph of stellar luminosity versus surface temperature. When stars are plotted on an H-R diagram, most lie along a band running diagonally across the graph from the upper left (luminous, hot, blue stars) to the lower right (faint, cool, red stars). This band is known as the "main sequence" because so many stars lie along it in the H-R diagram, and the stars along it are called "main sequence stars." Computer models of stellar interiors show that main sequence stars generate their energy by fusing hydrogen into helium in their cores. They are in a state of hydrostatic equilibrium, meaning the pressure due to gas and radiation trying to expand them is balanced by their self-gravity trying to contract them. Most stars spend most of their energy-producing lives on the main sequence.

Our solar system's Sun, for example, is a main sequence star. It is estimated to be about 4.6 billion years old. It has been on the main sequence, fusing hydrogen into helium in its core, for most of that time, all but the first few tens of millions of years of its existence. It has enough hydrogen in its core to continue doing so for approximately 5 billion more years.

Stars are born in interstellar clouds of gas and dust called nebulae. These nebula are composed mostly of hydrogen, with some helium and very small amounts of other elements. A typical nebula contains enough matter to form hundreds or thousands of stars. A portion of a nebula will begin to contract gravitationally. This contraction might be triggered by shock waves from a nearby supernova or by the nebula's passage through a spiral-arm density wave as the nebula orbits the center of its galaxy. As the gas contracts, half the energy released gravitationally goes to raising its temperature, and the other half is radiated away; it starts to shine as a protostar. The central part of the protostar heats the most, and collisions between the atoms there strip away their electrons, ionizing the gas. When the central temperature of the contracting protostar reaches a few million kelvins, it becomes hot enough to begin fusing hydrogen nuclei (the most abundant element in stars initially) into helium nuclei in its core. The initiation of hydrogen fusion in the core stops the gravitational contraction, and the protostar becomes a main sequence star.

Nuclear fusion reactions release energy because the total mass of the lighter nuclei that are fused together exceeds slightly the mass of the heavier nucleus that is produced. The small excess mass is converted into energy according to the famous Einstein relation between mass and energy, $E = mc^2$. In hydrogen fusion, which supplies the energy for main sequence stars, four hydrogen nuclei fuse into one helium nucleus.

For fusion to occur, the positively charged nuclei must overcome their electrostatic repulsion for each other; this requires high temperatures (so the nuclei are moving fast) and high densities (so the nuclei are reasonably close together). Only the cores of main sequence stars have high enough temperatures and densities to sustain fusion reactions, so that is where hydrogen fusion occurs and the energy is produced.

The energy flows from the core to the surface via two possible mechanisms: convection and radiation. Convection transports energy as rising bubbles of hot, low-density gas, while surrounding cooler, denser gas sinks. Radiation transports energy as a flow of photons of electromagnetic radiation. Radiation always occurs throughout the interiors of main sequence stars whether convection also occurs or not. Low-mass main sequence stars (about 0.1 times the Sun's mass) have convection occurring throughout the interior. Main sequence stars like the Sun have a radiative central region and a convective outer envelope. High-mass stars (about 10 times the Sun's mass) have a convective central region and a radiative outer envelope.

As a main sequence star consumes the hydrogen in its core, it becomes slightly brighter, although it remains near the main sequence. When the hydrogen in the core of a main sequence star is exhausted, the core contracts and heats up; basically the core is "searching" for a new nuclear fusion reaction to tap. Hydrogen fusion is transferred to a shell still rich in hydrogen surrounding the shrinking helium core. The outer part of the star expands, and as it does so, it cools off and becomes redder. (Notice the two simultaneous but opposite behaviors: the core contracts and heats up while the outer layers expand and cool off.) The star becomes larger and redder. It leaves the main sequence and becomes a red giant or red supergiant.

The length of time a star spends on the main sequence depends on its mass. Massive stars have more fuel, but they consume it much more quickly, so their main sequence lifetimes are short—perhaps as little as a few million years, compared to about 10 billion years for a star like the Sun. Low-mass stars have less fuel, but they consume it very slowly, so their main sequence lifetimes are long—perhaps as long as a trillion years. This is much longer than the age of the universe, which means that every low-mass star that ever formed is still a main sequence star.

APPLICATIONS

Most stars, including our Sun, are main sequence stars. Understanding the structure and stability of main sequence stars is important in many ways. The Sun has immediate effects on the solar system. Any changes in the Sun will produce changes on Earth. The Sun's energy output is directly responsible for Earth's weather and the sustenance of all life on Earth. A small change in the energy output from the Sun could result in serious consequences for Earth, such as radically altered weather patterns, climate change (global warming or cooling), and shifting agricultural areas. A constant radiative output by the Sun is so important that even a minuscule variation could ultimately affect Earth's geopolitical stability. There is evidence that the Sun's activity (and output) might vary over timescales of decades to centuries, and these variations may be linked to recorded changes in climate. When the Sun finally does leave the main sequence and become a red giant, it will be so bright that Earth's temperature will rise dramatically. Our oceans will boil away, our atmosphere will escape into space, the surface rocks will at least partly melt, and Earth will become uninhabitable.

Since the mid-1990's, planetary systems have been discovered around a number of other stars. Few of the planets detected so far seem at all Earth-like, but it is interesting to speculate about the possibility of extraterrestrial life and especially intelligent life. The only known example of the development and evolution of life (and intelligence) is on Earth, but if the timescale here is typical, then only main sequence stars of about one solar mass or less would have stable main sequence lifetimes long enough to allow life, and possibly intelligence, to develop on planets orbiting them. In a project called the Search for Extraterrestrial Intelligence (SETI), stars are monitored with radio telescopes to try to find modulated radio signals that might be coming from intelligent civilizations on planets orbiting those stars. The stars targeted for such monitoring for the most part are main sequence stars of one solar mass or less, because they are thought to be the best candidates to have planets on which life might have developed.

CONTEXT

The interior structure of stars cannot be observed directly, but instead is studied by constructing computer models, based on the physics that we think applies to different interior regions of different stars. Main sequence stars have comparatively simple interiors, and astrophysicists have a fairly good theoretical understanding of them. Consequently, main sequence stars provide an important link between the theoretical physics of stellar interiors and the stellar properties that are observed and measured directly.

One way of confirming theories about hydrogen fusion in main sequence stars generally and the Sun in particular is to try to detect the neutrinos that should be produced in the fusion reactions. Neutrinos do not interact readily with most matter, and so they should pass directly out of the Sun's interior. Beginning in the 1970's, sensitive neutrino detectors picked up only about one-third as many neutrinos as models of the Sun's interior predicted. This discrepancy was finally resolved in 2003, when the Sudbury Neutrino Observatory (SNO), deep underground in a nickel mine in Sudbury, Ontario, Canada, detected the predicted number of neutrinos, but of all three types and not just the one type produced in hydrogen fusion for which the early detectors were designed. This confirmed a hypothesis from particle physics that the three types of neutrinos could convert from one type to another and proved that neutrinos do have mass (albeit a very small mass), since only if they have mass can they change between types.

While addressing the astronomically interesting questions of stellar structure, the study of main sequence stars will continue to have direct application to our Sun's stability and its consequent effects on Earth's long-term climate and the welfare of humanity.

Dennis Chamberland

FURTHER READING

Bok, Bart J., and Priscilla F. Bok. *The Milky Way*. Cambridge, Mass.: Harvard University Press, 1981. The definitive book on the Milky Way written by two world experts on the subject. The book is easy to digest and includes a detailed deliberation of main sequence stars, the H-R diagram, and its relevance to the discussion of stars near the Sun. Illustrated.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Contains a thorough description, complete with transparent overlays, of the H-R diagram and the place of main sequence stars in stellar evolution.

Cooke, Donald A. *The Life and Death of Stars*. New York: Crown, 1985. This very colorfully illustrated and easy-to-read book gives an in-depth account of star birth, from nebula to protostar to main sequence. Provides an excellent discussion of stellar equilibrium states. For a wide audience.

Cornell, James, and Alan P. Lightman. *Revealing the Universe*. Cambridge, Mass.: MIT Press, 1982. This book discusses specific areas of observational and theoretical astronomy. A discussion of the H-R diagram is included, specifically on how the physics of stellar interiors determines the shape of the H-R diagram. Somewhat technical and best approached by those with a scientific background. Illustrated.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Has a good description of the H-R diagram and the place of main sequence stars in stellar evolution.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A thorough and well-written college-level introductory astronomy textbook, with a good description of the H-R diagram and the place of main sequence stars in stellar evolution.

Harwit, Martin. *Cosmic Discovery*. New York: Basic Books, 1981. Covers "the search, scope and heritage of astronomy," including the development of the H-R diagram and the recognition of main sequence stars. A fascinating expedition through the history of astronomy. Illustrated.

Kaler, James B. *Stars*. Scientific American Library Series 39. New York: W. H. Freeman, 1998. Covers all aspects of stars, including their origins, distribution, composition, types, and evolution. Illustrated and well written.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Divided into short sections on specific topics, this college textbook for introductory astronomy courses contains

a thorough discussion of the H-R diagram and the place of main sequence stars in stellar evolution.

See also: Brown Dwarfs; Gamma-Ray Bursters; Hertzsprung-Russell Diagram; Novae, Bursters, and X-ray Sources; Nuclear Synthesis in Stars; Planetary Formation; Protostars; Pulsars; Red Dwarf Stars; Red Giant Stars; Solar Evolution; Stellar Evolution; Supernovae; Thermonuclear Reactions in Stars; White and Black Dwarfs.

Mars: Illusions and Conspiracies

Categories: Mars; Planets and Planetology

Misinterpretations of telescopic observations led to the illusion of artificial canals on Mars, and visits of increasingly sophisticated spacecraft to the planet resulted in detailed photographs in which some enthusiasts saw faces and pyramids but space scientists saw buttes and mesas. These fanciful misperceptions can serve as a cautionary tale against the tendency by observers with strong preconceptions to see what they want to see rather than what is really there.

OVERVIEW

Some astronomers see the history of their discipline as the supersession of erroneous beliefs (or illusions) by true ideas that, over time, become ever more faithful descriptions of reality. In this way, the spherical Earth replaced a flat one, heliocentrism superseded geocentrism, and an expanding cosmos of multitudinous galaxies supplanted a static, small Milky Way universe. Outside Earth, Mars (the Romans' name for Ares, as it was known to the ancient Greeks) has been the subject of more illusions and conspiracies than any other planet. For centuries this "Red Planet" has fired the imagination of humankind, and the invention of the telescope and even visitations by orbiters and landers have not dampened the desires of many to see Mars as a locus of past or present intelligent life.

During the nineteenth century, ever more powerful telescopes allowed astronomers to see yellow clouds and white polar caps on the Martian surface. These astronomers commonly interpreted the polar caps, which expanded and diminished with the seasons, as frozen water. Others interpreted the dark areas as seas and the light areas as land. Pietro Angelo Secchi, a Jesuit astronomer who was the first to make multicolored drawings of Mars, was also the first to use the Italian term *canale* for a prominent feature (later called Syrtis Major). Another Italian astronomer, Giovanni Schiaparelli, created, starting in 1877, extremely detailed drawings of the Martian surface with hundreds of named features, including a growing number of rectilinear *canali*. Although *canale* may mean "canal," as in Canale di Panama, Schiaparelli interpreted these Martian *canali* as "channels" or "grooves," that is, as natural, not artificial, structures. He viewed these interlocking channels as a "hydrographic system" through which melting polar waters circulated throughout the planet, helping to foster organic life. He was "absolutely certain" that he had seen this intricate network, and he was even open to the possibility that the *canali* might be artificial, but he dismissed as imaginary the convoluted maps and interpretations of observers such as Percival Lowell.

For years the American businessman and amateur astronomer Lowell had successfully managed the family's great fortune, based on textiles, finance, and land, but, having read about Schiaparelli's *canali*, he became passionately interested in Mars and used his wealth to build an observatory near Flagstaff, Arizona, whose altitude and dry desert air facilitated Lowell's ability to see many more *canali* than Schiaparelli had detected. Lowell also became convinced that the complex system of hundreds of canals that he mapped was due to the constructive skills of intelligent Martians. He even invented a story to explain how these massive canals, which needed to be many kilometers wide to be visible from Earth, were built to bring polar water to the warm equatorial regions to support life on a planet growing colder and

Mars Compared with Earth

Parameter	*Mars*	*Earth*
Mass (10^{24} kg)	0.64185	5.9742
Volume (10^{10} km^3)	16.318	108.321
Equatorial radius (km)	3,396	6,378.1
Ellipticity (oblateness)	0.00648	0.00335
Mean density (kg/m^3)	3,933	5,515
Surface gravity (m/s^2)	3.71	9.80
Surface temperature (Celsius)	–128 to +24	–88 to +48
Satellites	2	1
Mean distance from Sun millions of km (miles)	228 (141)	150 (93)
Rotational period (hrs)	24.63	23.93
Orbital period	687 days	365.25 days

Source: National Space Science Data Center, NASA/Goddard Space Flight Center.

drier. He published his data, maps, and interpretations in three books, *Mars* (1895), *Mars and Its Canals* (1906), and *Mars as the Abode of Life* (1908), which were popular with the public, journalists, and science-fiction writers but were criticized by many astronomers and scientists. In the twentieth century, astronomers, who had the advantage of powerful refracting and reflecting telescopes, were unable to see either Schiaparellian or Lowellian canals. In 1907, Alfred Russel Wallace, who had discovered natural selection independently of Charles Darwin, published a book that attacked Lowell's hypothesis of a Mars inhabited by intelligent beings by showing that the planet was so cold and dry that it was "absolutely uninhabitable." Although a few astronomers such as Earl C. Slipher came to Lowell's defense, many were skeptical, since they failed to find any water vapor in the Martian atmosphere and since those who saw the canals failed to agree on their locations or dimensions. Nevertheless, Slipher, who took 126,000 photographs of Mars at Lowell Observatory from 1906 to 1962, insisted that many of these pictures verified the existence of linear canals.

Illusions about Mars increased through publications such as Immanuel Velikovsky's *Worlds in Collision* (1950), in which he used astronomical data, biblical stories, and folk legends to argue that the Earth, Venus, and Mars had exchanged atmospheres when, only a few thousand years before, they nearly collided with each other. Even though the American scientific community vigorously attacked Velikovsky's ideas, which were easily falsified by scientific information and analysis, he developed a cult following. He and his followers attributed the derision of scientists to a conspiracy to cover up the real truth behind the solar system.

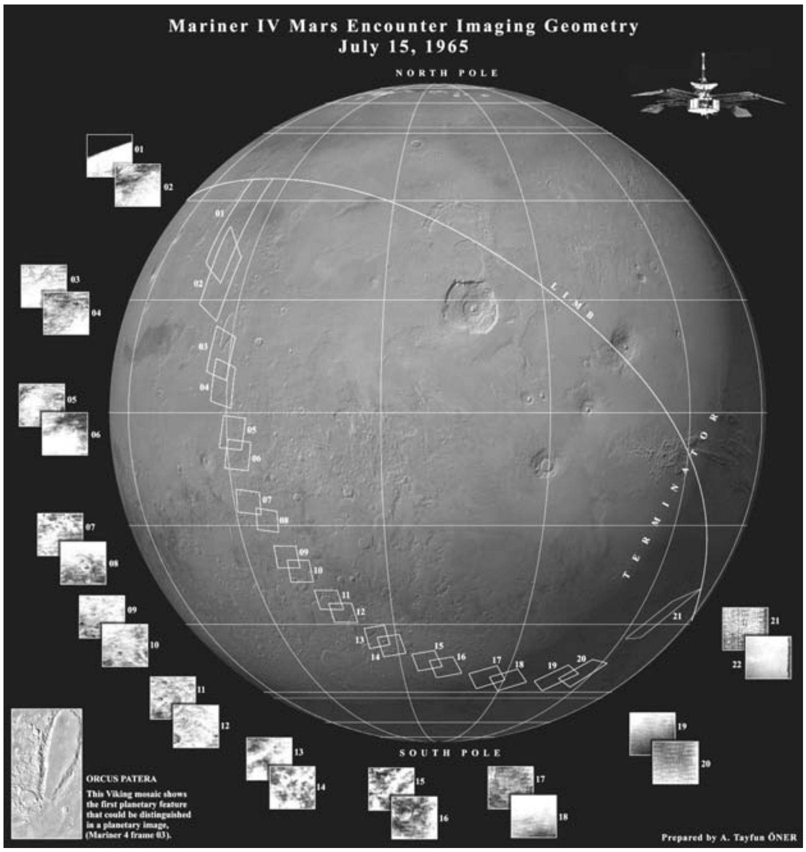
In 1965 the American spacecraft Mariner 4 became the first to send back pictures of the Martian surface from a relatively close range. These pictures confounded the believers in canals, Mars cultists, the public, and astronomers, who were surprised by an extensively cratered world that was more Moon-like than Earth-like. Furthermore, Mars's extremely thin atmosphere, only one two-hundredth as extensive as Earth's atmosphere, was unable to support any liquid water. Other Martian probes, such as Mariners 6 and 7, appeared to confirm the view of Mars as lifeless and dull, far from the exciting vision of Lowell and his followers.

KNOWLEDGE GAINED

The conception of what some astronomers called the "old Mars" was transformed by Mariner 9, which arrived at the planet late in 1971 and comprehensively mapped its surface in 1972. Its dramatic images revealed a new and more interesting Mars, with such fascinating features as Olympus Mons, the solar system's largest volcano, and Valles Marineris, a complex of ravines so vast that it could hold nearly five hundred of the Earth's Grand Canyon. Mariner 9's many detailed pictures of mountains, craters, mesas, plains, valleys, and other features provided areologists—those who study Mars—with an abundance of data to construct a geological history of Mars. These images also, however, encouraged speculators to develop controversial interpretations of ambiguous features. These speculations became highly imaginative when sharper pictures of Mars were returned to Earth by Viking Orbiter 1, especially of Cydonia, a plain with scattered buttes and mesas in the middle northern latitudes. One photograph, taken on July 25, 1976, showed a mesa that, in certain light, vaguely resembled a human face, and even though space scientists saw it as natural, similar to anthropomorphic geological features on Earth, others interpreted it as a remnant of a vanished civilization.

When, in 1979, computer scientists released enhanced images of the Cydonian region, some interpreters saw such new features in the “face on Mars” as eyes with pupils, a nose, a mouth with teeth, and a headdress reminiscent of an Egyptian pharaoh’s. These interpreters also claimed to see several pyramids in the region surrounding the “face.” These pyramids were hundreds of times larger than the Great Pyramids at Giza in Egypt, and, according to these interpreters, their geometrical shapes, which were a clever combination of pentagonal and hexagonal symmetries, as well as their architectural alignments, manifested a knowledge of astronomical and mathematical constants.

Magazine articles, books, and Web sites soon proliferated, advancing the hypothesis that the Cydonian region contained the remains of a colossal city. Richard Hoagland, a journalist who had become interested in the Cydonian structures in 1980, published a book, *The Monuments of Mars* (1987), which became so popular that it went through four editions in the late 1980’s and into the 1990’s. When space scientists ridiculed the interpretations of the “Cydonian cultists” by showing that their “pyramids” were actually natural structures that had been sculpted by wind and water, not intelligent beings, the “Cydonians” responded by founding such organizations as the Academy of Future Science and the Independent Mars Investigation Project to defend their views. Using publications, Web sites, radio and television appearances, and DVDs, these Cydonians alleged that a conspiracy of the National Aeronautics and Space Administration (NASA) and other governmental agencies had developed to cover up the evidence of intelligent life on Mars, since this explanation challenged established social, political, and religious values. When such missions as the Mars Observer failed, some Cydonians accused NASA scientists of deliberately sabotaging the mission to prevent the public from learning “the truth” about the “face and pyramids.” When such



In 1965, Mariner 4 was the first probe to reach Mars. It supplied data and images that initiated our modern understanding of the Red Planet: a cratered, harsh surface rather than a lush environment capable of supporting life. (NASA)

NASA missions as the Mars Global Surveyor, Mars Pathfinder, and Odyssey succeeded, returning data that showed the "face" as an eroded conical butte and the "pyramids" as irregularly shaped mesas, the Cydonians reprimanded the scientists for interpreting these new photographs through their academic prejudices. Some critics wondered whether the evidence gathered by a lander or astronauts in the region of the supposed face and pyramids would be sufficient to resolve the conflicts between space scientists and the Cydonians.

CONTEXT

Throughout history, genuine scientists have often contended with pseudoscientists. Sometimes, a pseudoscience has evolved into a genuine science (as in the case of alchemy, which became chemistry), but astronomers have been unable to dispatch pseudoastronomies such as astrology. As planetary scientist Carl Sagan often remarked, even in such advanced societies as the United States, more professional astrologers were making a living than professional astronomers. Many space scientists considered any attention paid to pseudoscientific theories about Mars as a waste of their time and taxpayers' money. For example, instead of spending valuable time, money, and human energy in fulfilling the assigned mission of the Mars Global Surveyor, scientists, because of pressure from the media, public, and politicians, were forced to divert the spacecraft to take pictures of the "face on Mars" and other supposed "alien constructs." These new pictures clearly revealed, in the view of the scientists, that the objects in question were naturfacts, not artifacts.

Other scholars have studied these illusions and conspiracies for the insights that they provide into fascinating psychological and sociological phenomena. For example, the cases of Schiaparelli and Lowell are illustrative of "pathological science," in which scientists erroneously convince themselves of the genuineness of what they are seeing and saying. Pathological science often occurs when threshold effects are involved—for example, when scientists attempt to interpret things that are barely visible. In these cases humans tend to overinterpret scattered dark features into a line, as Schiaparelli and Lowell did. In the case of the Cydonian cultists, other forms of misinterpretation have surfaced. For example, modern computerized image enhancement can be overdone, sometimes leading to the appearance of features that were not present in the original. Sociological pressures often help explain how Cydonians see themselves as the victims of various conspiracies. They find pleasure and comfort in the idea that intelligent creatures once inhabited Mars, but these extraordinary claims require extraordinary proofs. However, scientists, relying on data collected by superior orbiters and landers, have seen nothing that cannot be better understood as the consequence of natural forces. History has shown that humans can fool themselves, but scientists, based on long experience, have found that "Mother Nature" cannot be fooled.

Robert J. Paradowski

FURTHER READING

Goldsmith, Donald, ed. *Scientists Confront Velikovsky*. Ithaca, N.Y.: Cornell University Press, 1977. Eminent scientists such as Carl Sagan show why, from a detailed study of astronomical evidence, Velikovsky's theories about the solar system in general, and Mars in particular, must be regarded as pseudoscience. Bibliography, index, and references for each chapter.

Hartmann, William K. *A Traveler's Guide to Mars: The Mysterious Landscape of the Red Planet*. New York: Workman, 2003. Called a "masterpiece of scientific writing for the general reader," this lavishly illustrated book provides a fascinating survey of the "New Mars," including sections and sidebars on how astronomers have responded to illusions about Mars. Glossary, selected sources, and index.

Hoyt, William G. *Lowell and Mars*. Tucson: University of Arizona Press, 1976. A critical but balanced portrait of Lowell, his observatory, his writings, and the influence of his ideas on the public, with an account of the discovery of Pluto, named in his honor.

Kargel, Jeffrey S. *Mars: A Warmer, Wetter Planet*. New York: Springer Praxis, 2004. A detailed analysis of the new face of Mars that has been revealed by the evidence gathered by orbiters and landers. Illustrated with a color section and many black-and-white photographs. Bibliography and index.

Kieffer, Hugh H., et al., eds. *Mars*. Tucson: University of Arizona Press, 1992. This massive compendium, part of the Space Science series, is the product of 114 collaborating experts. Part 1 deals with the history of explorations of Mars visually, telescopically, and through spacecraft. Here the subject of illusions about Mars is also treated. Has an excellent section, "Books About Mars," on pages 19-24. Illustrated with photographs and diagrams. Glossary, bibliography, and index.

Sheehan, William, and Stephen James O'Meara. *Mars: The Lure of the Red Planet*. Amherst, N.Y.: Prometheus Books, 2001. Sheehan, a psychiatrist and amateur astronomer, analyzes the history and controversies connected with human encounters with Mars through the naked eye, telescopes, and modern spacecraft. Intended for general audiences, the book has appendixes offering data on Mars, notes to original and secondary sources, and an index.

Wilford, John Noble. *Mars Beckons: The Mysteries, the Challenges, the Expectations of Our Next Great Adventure in Space*. New York: Alfred Knopf, 1990. Wilford, a two-time Pulitzer Prize winner, discusses the history of both illusions about and hard-won knowledge of Mars, with a concluding treatment of future plans for getting humans on the planet. Appendixes on the Mars missions and on the solar system. Bibliography and index.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Mars: Possible Life; Mars: Surface Experiments; Mars's Atmosphere; Mars's Craters; Mars's Polar Caps; Mars's Satellites; Mars's Valleys; Mars's Volcanoes; Mars's Water; Meteoroids from the Moon and Mars; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits; Planetary Rotation; Planetary Satellites; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets.

Mars: Possible Life

Categories: Life in the Solar System; Mars; Planets and Planetology

Environmental conditions in the past suited the origin of life on Mars. A primary motive for sending orbiters and landers to the planet has been to detect Martian life; it would be a principal goal for a crewed mission as well. If detected, such life would help elucidate the origin of life on Earth and, possibly, elsewhere in the universe. If not detected, the question of whether terrestrial life is unique would remain open.

OVERVIEW

On the basis of their observations from telescopes, photographs and remote sensors from orbiters, and direct inspection by means of experiments performed by landers and rovers on the planet's surface, scientists have ruled out the presence of intelligent life on Mars, fourth planet from the Sun and Earth's immediate outer neighbor. No direct evidence of life of any kind has been found. However, circumstantial evidence that life might have existed in the past, and could persist even today, steadily accumulated after the first probe photographed Mars in 1964. In all cases, it is evidence based on comparisons with the conditions that support life on Earth.

Mars specialists base their search on the broad biological definition of life as a chemical system capable of Darwinian evolution to accommodate to changing environmental conditions. Most agree that this entails the ability to process energy and nutrients, grow, and reproduce. Accordingly, the search is on for environments on Mars that could foster these activities. Unfortunately, it is a harsh world. The Viking landers of the mid-1970's confirmed what scientists had suspected, that iron oxide, poisonous to life on Earth, permeates the planet's surface. The atmosphere is thin, frigid, dry, and very low in oxygen. The absence of a planetary magnetic field permits intense ultraviolet radiation, also deadly to life, to reach the surface. These facts suggest that the existence of life above ground is highly unlikely.

Water is an essential element for life, although its presence does not guarantee life. According to evidence from orbiters and landers, Mars had oceans, lakes, rivers, and a thicker atmosphere in the past, and may still see occasional outbursts of underground water on the surface. However, most water now exists as ice in polar ice fields and as permafrost beneath the surface, a region known as the cryosphere. Below these there may be a hydrosphere, a band of water-permeated rock. Most scientists believe that life could thrive at the boundary between the cryosphere and hydrosphere. There is some evidence to encourage that belief. Probes have detected methane in Mars's atmosphere; methane is a waste gas from biological processes, and it is known as a biomarker or biosignature. Because methane lasts in the atmosphere for only a few hundred years, it must be replenished. That source could be Martian organisms now alive, although theoretical inorganic chemical processes also have been proposed as the source.

If organisms exist in the cryosphere-hydrosphere boundary region or as spores near ancient water bodies, they are probably not large. There could be multicellular organisms like the tube worms that feed from hydrothermal vents in the cold dark waters of the deep oceans on Earth, but most scientists foresee finding only single-cell organisms. Terrestrial organisms known as extremophiles live in conditions ranging from 253 to 394 kelvins (–4° to 250° Fahrenheit), in a wide range of acidity, in very salty water, without light, or kilometers underground. Some extract energy and nutrients from hydrothermal vents (hyperthermophiles), such as hot springs. Others feed from inorganic chemicals in rocks (chemolithoautotrophs)—for instance, ingesting sulfide minerals and excreting sulfuric acid—and in slushy water or salty water colder than the normal freezing point (psychrophiles). It is Martian equivalents of such extremophiles that scientists hope to find.

In a subsurface ecosystem Martian organisms would show variety in form and function. The basic structure ought to be a cell with a semipermeable membrane, such as the sack of lipoproteins defining most terrestrial cells, and internal structures that split apart chemicals in order to use the by-products in their metabolism. Most would be grazers, feeding off the ambient nutrient source, but there are likely to be predators that consume the grazers. Either may have a means of locomotion, such as cilia or the ability to expand and contract. In order to evolve, they would need some type of chemical record of their mechanism, as exists in the ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) of Earth organisms, to pass on during reproduction. Because the temperature and air pressure on Mars are low, the metabolism, reproduction, and movement of organisms may well be sluggish and sparse in comparison to those functions in Earth organisms. They could live from as little as 35 centimeters (1.1 foot) below the surface to as much as 10 kilometers (6.2 miles), scientists speculate.

Martian life may have once thrived and then declined and finally vanished as the ancient surface water dried up. In that case, fossils might remain. Paleobiologists have found fossilized forms of Earth's single-cell organisms, as in stromatolites, that date back at least 3.5 billion years before the present. Should such fossils survive on Mars, they would establish that life rose there and provide clues about which environmental conditions supported an ecosystem. In 1996 scientists from the National Aeronautics and Space Administration (NASA) announced finding evidence for life in a Martian meteorite known as ALH 84001, discovered in Antarctica. Among the evidence was what appeared to be a fossilized cellular structure. Although the majority of scientists later discounted the evidence and judged the microscopic structure to be the result of nonbiological processes, research demonstrates that the conditions for preserving delicate structures, such as fossils, exist on Mars. Additionally, it is possible that early conditions were favorable for life but then, because of meteorite bombardments 3.9 billion years ago, changed before life actually got started. In that case, scientists might still detect the precursors to life in complex organic chemicals.

Two further possibilities for life on Mars worry scientists. Just as ALH 84001 was blasted off Mars by a meteorite and made its long journey to Earth, a chunks of Earth probably reached Mars, and they could have carried Earth organisms there, seeding the Red Planet with terrestrial life. It is probable, moreover, that the landers sent to Mars by American, Russian, and European space agencies carried organisms with them despite decontamination protocols. Such “forward contamination” would be even more likely from a crewed mission to the planet. Seeding and contamination do not mean that Earth organisms have survived in the harsh Martian environment. On the other hand, if Earth organisms have adapted, scientists may have difficulty establishing their origin.

KNOWLEDGE GAINED

Experiments performed aboard the Viking landers (1976) established that Mars’s soil contains chemicals inimical to life, particularly oxidants such as iron oxide (rust), and the Sojourner, Opportunity, and Spirit rovers confirmed the finding at other locations. These findings do not absolutely rule out life. Viking experiments may have missed a biomarker that they were not designed to detect, or the rocket exhaust from their landing may have killed organisms within their reach. In any case, the search for life taught scientists much about the chemical nature of the soil and the distributions of oxidants.

In the 1990’s NASA set a policy for its Mars probes: “Follow the water.” The resulting search from orbit and on the surface found evidence of erosion and chemical deposits, such as hematite, that on Earth derive from flowing water. Sensors in orbiters detected underground ice as well. These discoveries not only encourage scientists to look further for life but also suggest that sufficient water exists to support human habitation on the planet. At the same time, the various landers and rovers, all of which far exceeded their expected performance, proved the versatility and hardiness of technology in the Martian environment.

When it became probable that no life inhabited the Martian surface, scientists investigated the possibility of organisms living below the surface. The research depended on analogy to similar Earth habitats, and scientists were inspired to search for life in heretofore unexplored realms. The result was to expand greatly the knowledge about terrestrial extremophiles near volcanic vents deep in the oceans, in gelid water, in porous underground rock, or in salty, acidic, or alkaline conditions. Paleobiologists uncovered fossilized organisms from much further in the past than previously suspected.

In late 2008, analysis of data from NASA’s Mars Reconnaissance Orbiter (MRO)—reported just after that spacecraft was given a two-year extension beyond completion of its primary mission—indicated that Mars in its distant past must have had sufficient water flowing across the surface that clay-rich (carbonate) minerals could form. MRO picked up the signatures of those clay minerals. That Mars may have been wetter and favorable for primitive life to develop is evidenced by the fact that these clay minerals survived to the present. As Mars began losing its water, becoming a drier planet, remaining water would have become acidic. Carbonate clay minerals are relatively easily dissolved in acidic water. Where the clay minerals survived on the surface to the present era would have been locations less hostile to life as the planet continued to become drier and drier. Although highly suggestive, this new information does not provide direct evidence that life may ever have existed on Mars.

Another discovery, this one made by ground-based telescopes atop Mauna Kea, Hawaii, was reported in an early January, 2009, *Science Update* aired on NASA Television. Reporting scientists explained that the infrared signature of methane gas had been detected in significant amounts in the Martian atmosphere. This held the potential for indirect evidence that primitive life might exist on the Red Planet in the present era. Since solar ultraviolet radiation penetrates through the thin Martian atmosphere, methane molecules dissociate, and over the eons of geologic time the methane should have dissipated without replenishment. Replenishment was possible by either one of two mechanisms: a geological mechanism involving the conversion of iron oxide to serpentine minerals in the presence of water, carbon dioxide, and a heat source, or a biological mechanism involving digestion occurring in primitive microorganisms. More direct examination of methane vents on the planet’s surface would be needed to determine if methane production is of geological or biological origin.

Negative findings are also important, as much to human culture as to science: Mars supports no civilization or, in all probability, animal life. The alteration in surface color through the Martian seasons that fascinated astronomers like Percival Lowell comes from wind storms, not vegetation. Thereby, scientists discounted the possibility, so popular in science fiction, that Martian life poses a threat to Earthlings.

CONTEXT

The absence of Martians as imagined by H. G. Wells, Robert A. Heinlein, or Ray Bradbury has not dampened popular enthusiasm for the search for life on Mars. It endures because it promises to answer a question that has long made humanity look to the heavens and wonder: "Are we alone?"

The discovery of life on Mars would have profound implications. If proven to be entirely independent of life on Earth, that finding would mark a shift in understanding the universe as great as that of the Copernican Revolution. Philosophy would be tasked to reconsider the human moral obligation to other organisms. Religions would have to cope with the fact that life on Earth is not a unique creation. Science would be encouraged to look for life on still more worlds, such as the satellites of Jupiter, and for biomarkers in the light from other planetary systems. At the same time, NASA and other space agencies would need to take measures to protect Martian life from Earth organisms hitchhiking aboard planetary probes, or, in the case of a sample-return mission or crewed mission, to protect Earth's ecosystem from Martian organisms.

If Martian organisms were found to be related to those on Earth, the knowledge would also be fundamentally important. It would establish the great durability of life in spreading from one planet to another and leave open the possibility, as proposed in the Panspermia hypothesis, that both Mars and Earth were seeded long ago with life that originated elsewhere.

Proof that life never existed on Mars would be significant as well. The question of Earth's uniqueness would remain unsettled, yet scientists would learn an important fact: The types of chemical and geophysical conditions found on Mars are not conducive to life. Why that should be true would pose a major question for further research.

Roger Smith

FURTHER READING

Elkins-Tanton, Linda T. *Mars*. New York: Chelsea House, 2006. In a style suitable for high school students, the author, a geophysicist, provides a technically specific overview of Mars's environment and geologic history, including a short, reasonable summary of the case for life. With glossary, appendixes on basic relevant science, and photos.

Forget, Françoise, Françoise Costard, and Philippe Lognonné. *Planet Mars: Story of Another World*. Chichester, England: Praxis, 2008. The lucid text, suitable for high school students, explains current knowledge about the geology, hydrology, and meteorology of Mars. Three short chapters discuss possible life, and a fourth summarizes knowledge gained by space probes. With abundant photographs and graphics.

Hartmann, William K. *A Traveler's Guide to Mars: The Mysterious Landscapes of the Red Planet*. New York: Workman, 2003. This entertaining overview of the Red Planet includes a section about the physical conditions on Mars that could support life, and the text is a model of popular exposition. With many photographs and illustrations by the author, a planetary scientist, writer, and artist.

Kargel, Jeffrey S. *Mars: A Warmer, Wetter Planet*. New York: Springer Praxis, 2004. A planetary scientist, Kargel provides a comprehensive, sophisticated survey of the geology and climate of Mars throughout its history. The importance of water is a central theme, and he discusses the possibilities for life. With many graphics and photographs.

Kiang, Nancy Y. "The Color of Plants on Other Worlds." *Scientific American* 298 (April, 2008): 48-55. Although not specifically addressing possible Mars organisms, this article explains the nature and type of chemical evidence that scientists think will signal the presence of life in alien environments.

Walter, Malcolm. *The Search for Life on Mars*. Cambridge, Mass.: Perseus Books, 1999. Charming, concise, and equitable, this volume is especially valuable for Walter's explanations of how terrestrial paleobiology provides clues of what life on Mars might be like and the controversies and personalities behind the research.

See also: Extraterrestrial Life in the Solar System; Habitable Zones; Life's Origins; Mars: Illusions and Conspiracies; Mars: Surface Experiments; Mars's Atmosphere; Mars's Craters; Mars's Polar Caps; Mars's Satellites; Mars's Valleys; Mars's Volcanoes; Mars's Water; Search for Extraterrestrial Intelligence.

Mars: Surface Experiments

Categories: Mars; Planets and Planetology

The surface of Mars, as seen in telescopic observations and spacecraft images, hints at a history of past liquid water and possible habitability for life. Experiments conducted by robotic landers and rovers have helped scientists better understand the rocks and dust that form the planet's surface, as well as the role that water may have played in forming these materials.

OVERVIEW

While the surface of Mars has been studied extensively by telescope and spacecraft imaging (from both flybys and orbit), scientists' understanding of the surface materials and climate has been improved by data collected by robotic spacecraft located on the surface. The National Aeronautics and Space Administration (NASA) has mounted several missions, including the Viking, Mars Pathfinder, and Mars Exploration Rover (MER) missions, that have provided a context for orbital and telescopic observations and has advanced our insights into the geologic history of Mars.

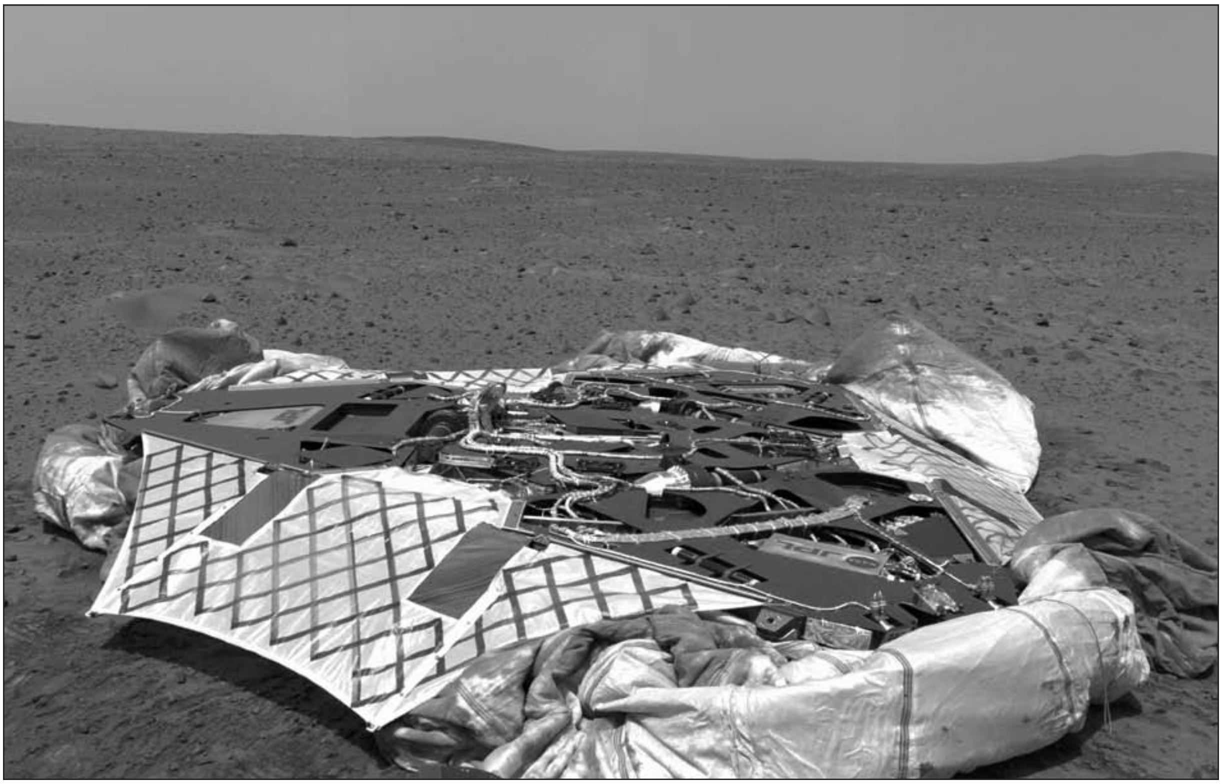
The Viking mission consisted of two identical spacecraft, each with an orbiter and lander. They were launched in 1975 and arrived at Mars almost a year later. Both landers touched down in the smooth northern plains: Viking 1 in Chryse Planitia (22.3° north, 48.2° west) and Viking 2 in Utopia Planitia (47.7° north, 225.9° west). The landers contained a suite of instruments for studying the environment, a 3-meter-long robotic arm to collect surface samples for analysis, and two cameras. Images could be acquired in visible or near-infrared wavelengths in 360° around the spacecraft from 40° below to 60° above horizontal. The elemental composition of surface materials was measured using an X-ray fluorescence spectrometer and atmospheric samples were analyzed using a gas chromatograph/mass spectrometer. The latter instrument was designed to search for organic materials as well. Further tests for evidence of life included the pyrolitic release experiment (to identify metabolism of carbon dioxide and carbon monoxide), the labeled release experiment (to detect metabolism of nutrients), and the gas exchange experiment (to examine gases released when nutrients and water vapor were added to a sample). Additional instruments included magnets (to identify magnetic components of samples), seismometers, and meteorological stations that measured wind speed and temperature.

Mars Pathfinder, launched in December, 1996, bounced onto the surface of Mars (cushioned by airbags) on July 4, 1997. Its landing site (19.3° north, 33.6° west)—which was 800 kilometers east-southeast of the Viking 1 site—was selected for safety and as part of NASA's "follow the water" strategy for Mars exploration. This location is at the mouth of Ares Vallis, a 1,500-kilometer-long outflow channel likely formed by catastrophic flooding. Rocks deposited there could have originated in any geologic unit crossed by the channel, so a variety of rocks and geologic time periods would likely be present (although source regions for the rocks would not be identifiable). The results were expected to increase knowledge of the geology and geochemistry of the Martian surface.

The Pathfinder mission consisted of a lander (named the Carl Sagan Memorial Station) and a rover (Sojourner). The Imager for Mars Pathfinder (IMP), mounted on the lander's 1-meter-high mast, provided panoramic stereo and multispectral imaging (up to eight wavelength bands) for rover navigation and science analysis. A set of magnets were also part of the IMP: images of the magnets were acquired throughout the mission to determine how much magnetic material from the Martian dust had been attracted to them. The final instrument system on the lander was the Atmospheric Structure Instrument/Meteorology package (ASI/MET), which measured atmospheric structure and temperature during descent and meteorological conditions (including temperature, wind speed, and wind direction) during operations.

The Sojourner rover, measuring 65 by 48 by 30 centimeters, was capable of moving up to 500 meters from the lander. The rover used cameras (two front-mounted stereo grayscale cameras, and one color camera on the rear) for navigation and science analysis. The rear camera also provided context images for chemical analyses by the Alpha Proton X-ray Spectrometer (APXS), also mounted on the rear. This instrument provided elemental compositions of rock targets (excluding hydrogen and helium) by bombarding the rock with alpha particles and measuring what type and energy of particle (alpha particle, proton, or X-ray photon) was backscattered off of the rock.

The Mars Exploration Rovers—MER A (Spirit) and MER B (Opportunity)—were launched separately in June, 2003, and landed on Mars in January, 2004. Landing sites were chosen based on engineering constraints, safety, and a continued scientific interest in evidence for past surface water. Spirit landed on the floor of Gusev crater (14.7° south, 184.5° west), a 170-kilometer-diameter impact crater that is bordered on the north by the volcano Apollinaris Patera. Gusev's southern rim is breached by a valley (Ma'adim Vallis) that ter-



A mosaic image compiled from the panoramic camera aboard the Mars Exploration Rover Spirit reveals the landscape around the rover's landing site at Gusev Crater, Mars, on January 18, 2004. (NASA/JPL/Cornell University)

In the mid-1970's, the Viking 1 and 2 landers conducted the first long-term explorations of the Martian surface; this image was captured by Viking 2 on the Utopian Plain. (NASA)

minates in the crater. The water that carved the valley may have once formed a lake in the crater. The landing site for Opportunity (2° south, 5.9° west) is located in Meridiani Planum, which was identified in orbital data as containing relatively high concentrations of the mineral hematite (Fe_2O_3). On Earth, such concentrations are often formed by deposition in standing bodies of water.

The two MER rovers are identical in design. Each rover utilizes multiple imaging systems, including the Panoramic Cameras (PANCAMs), used for navigation and hazard avoidance, and a Miniature Thermal Emission Spectrometer (Mini-TES). The PANCAM and Mini-TES optics are mounted on the 1.5-meter-high rover mast. PANCAM provides high-resolution images in fourteen wavelengths (visible and infrared) as well as stereo coverage in any orientation around the rover (360° panorama with 180° elevation coverage). These images are used for traverse planning and scientific analysis of both surface features and atmospheric conditions. Mini-TES detects thermal infrared radiation emitted by surfaces. Because different minerals emit different amounts of this radiation, Mini-TES can be used both to measure the temperature of a surface and to identify the minerals present.

Each rover has a robotic arm that can extend to reach rock targets up to 0.8 meters away. The arm is equipped with a Rock Abrasion Tool

(RAT) to brush away dust or expose unweathered interiors of rocks. Scientific targets are examined using the Microscopic Imager (MI). This instrument acquires images of a 31-millimeter-square area with a resolution of 30 microns per pixel. Compositional analyses are peformed using a Mössbauer spectrometer, which is sensitive to minerals containing iron, and an APXS instrument similar to the one on the Mars Pathfinder.

KNOWLEDGE GAINED

The Viking landers painted a bleak picture of the Martian surface, recording temperatures from 183 to 238 kelvins (–130° to –31° Fahrenheit) and high amounts of ultraviolet radiation. Images of the landing sites show a dusty, rock-strewn surface, with seasonal formation of extremely thin water-ice frosts at the Viking 2 site. Chemical analyses of the bright red surface sediments are most consistent with iron-rich clay minerals and contain relatively high amounts of sulfur and chlorine, possibly concentrated by evaporation of water. In places, this sediment forms a thin (1- to 2-centimeter), hard crust called a duricrust. Rocks at these sites appear dark in color, have a vesicular texture, and are likely basalt. The biological experiments proved ambiguous, as the observed results can be explained by abiotic chemical reactions. Both landers functioned for several years: Viking 1 until 1982, Viking 2 until 1980.

Mars Pathfinder images show a similar terrain, with red dust blanketing randomly strewn dark rocks, although some rocks appear to be oriented along the direction of flooding. Compositional analyses, however, suggested more than basalt. Several Pathfinder results are consistent with basaltic andesite, a more silica-rich volcanic rock. Formation of basaltic andesite requires recycling of basaltic crust, melting (and then erupting) the lighter minerals and leaving the heavier behind. On Earth, this process involves water-rich sediment and plate tectonics; the formation process for basaltic andesites on Mars is unknown. Pathfinder also provided detailed evidence of aeolian (wind) processes on Mars during its 83-sol mission, returning images of dust devils, wind streaks, and dunes.

The foot of the Mars Phoenix Lander and surrounding Martian surface at the north polar region, taken on May 25, 2008. (NASA/JPL-Caltech/University of Arizona)

The Mars Exploration Rovers have significantly changed our understanding of the geology of Mars. While initial analyses by Spirit found olivine-bearing basalts along the floor of the Gusev crater (unaltered volcanic rocks, not the expected lakebed deposits), areas studied later in the Columbia Hills, 3 kilometers from the landing site, included silica-rich sediment (possibly from an ancient hydrothermal system), iron sulfates, and evidence of explosive volcanism (which requires more silica-rich magma than basaltic eruptions, requiring crustal recycling).

Martian soil on the doors of the Thermal and Evolved Gas Analyzer (TEGA), onboard the Phoenix Lander, October 7, 2008. (NASA/JPL-Caltech/University of Arizona/Max Planck Institute)

Opportunity provided the most conclusive evidence for the presence of past surface water on Mars. Located fortuitously in a small crater, the first images showed an expanse of rock outcrop, not the scattered rocks seen at other sites. This outcrop and others examined later contain fine layers common in sedimentary rocks; some of these layers are inclined, forming crossbeds that indicate ripples or dunes. Chemical analysis identified light-colored rock in this outcrop as jarosite, an iron sulfate mineral that forms by evaporation of water. Embedded in this layer are spherules of gray hematite (dubbed "blueberries" by the science team), similar to concretions of iron oxide that form in aqueous environments on Earth.

CONTEXT

Mars has long piqued the curiosity of observers with hints of past liquid surface water. Early surface experiments suggested a surface dominated by volcanic rock (basalt), while more recent results point to a more complex geologic history with possible bodies of water, hydrothermal systems, and recycling of the basaltic crust. The Mars Exploration Rovers continue to return data that paint a more detailed geologic history of their local areas of Mars, and the Mars Phoenix lander in 2008 examined the possibility of the existence of subsurface ice and habitats for life in the Martian arctic (near 68° north). Before the arrival of winter at the Mars Phoenix landing site (near the end of calendar year 2008), a gas analysis of samples taken from the subsurface confirmed what photographs of trenches dug by the lander's robotic arm had strongly suggested, that being that indeed just centimeters below the surface water ice was present. When exposed to ambient conditions, that water ice sublimated into the Martian atmosphere. As a result of ongoing Mars exploration by robotic spacecraft in orbit and on the surface, our understanding of the Martian surface is rapidly changing.

In early 2009, NASA aired a special *Science Update* program that revealed that astronomers using a pair of telescope facilities high atop Mauna Kea in Hawaii had detected the infrared spectral signature of methane in the atmosphere of Mars. Given the nature and thinness of the Martian atmosphere, atmospheric methane could not survive the constant irradiation from solar ultraviolet rays. As a result, the implication was that this methane, found in relatively significant amounts, had to be replenished. Two production methods were possible, one geological and the other biological. On Earth, methane can be produced geologically in the conversion of iron oxide to serpentine minerals. This requires carbon dioxide, water, and an internal heat source. Mars has carbon dioxide, and thanks to the results from Mars Phoenix in 2008, it is known to have subsurface water at least in some locations. With a source of heat, serpentine mineral production on Mars might be possible. A biological origin would involve digestive processes by microorganisms. Data presented at this *Science Update* could not be used to discern whether biology or geology was responsible for the methane.

At the time of this important announcement, NASA was well into the selection process for identifying suitable landing sites for the Mars Science Laboratory (MSL), set for a 2011 launch. This discovery had the potential to shift the three-decade-old paradigm used in Mars exploration: following the water. Instead, consideration was immediately given to the search for a methane vent on the surface of the Red Planet, so that the nuclear-powered MSL might be able to conduct in situ analyses of samples and thereby determine whether Mars's methane was of geological or biological origin. If the latter proved to be true, it was likely that the primitive life-forms would subsist in subsurface layers in the presence of water and heat.

Jennifer L. Piatek and David G. Fisher

FURTHER READING

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. Written by top scientists, these articles cover many aspects of solar system research. The Mars chapter contrasts the chemical analyses from the Viking and Mars Pathfinder missions.

Bell, Jim F., ed. *The Martian Surface*. Cambridge, England: Cambridge University Press, 2008. Compiles scientific results from 1992 to 2007 in a series of articles written by prominent Mars scientists. Provides a continuation of the last comprehensive review of Mars science by Kieffer et al. (see below). An excellent source of citations. Requires a strong background in geology, chemistry, and physics to grasp the concepts fully.

Boyce, Joseph M. *The Smithsonian Book of Mars*. Washington, D.C.: Smithsonian Institution Press, 2002. An accessible overview of Mars exploration, covering telescopic observations, mission results, the search for life on Mars, and future exploration. Places surface results from Viking and Pathfinder in context with orbital observations.

Carr, Michael H. *The Surface of Mars*. Cambridge, England: Cambridge University Press, 2005. A beautifully illustrated and citation-rich text that describes the current understanding of the Martian surface, discusses the geologic processes involved in its formation, and compares these processes to those on Earth. A similar summary for Viking results is presented in a 1981 book with the same title and author (Yale University Press).

Hartmann, William K. *A Traveler's Guide to Mars*. New York: Workman, 2003. Written in a light, engaging style, this book presents the exploration, geology, and geography of Mars in the format of a tourist's guidebook, including maps and images.

Kieffer, Hugh H., et al., eds. *Mars*. Tucson: University of Arizona Press, 1992. Review articles by top researchers provide the definitive summary of the post-Viking view of Mars. This weighty tome contains a wealth of information on nearly every topic in Mars research and is a good source for citations. A strong science background (geology, chemistry, and physics) is necessary to understand many of the concepts presented.

Raeburn, Paul, and Matt Golombek. *Uncovering the Secrets of the Red Planet*. Washington, D.C.: National Geographic Society, 1998. Written in an anecdotal style, this large-format book tells the story of Mars exploration, particularly Mars Pathfinder (author Golombek was the project scientist) from both science and engineering viewpoints. Contains many mission images, including three-dimensional anaglyphs (glasses are included) and panoramas, to illustrate the scientific results.

Sheehan, William, and Stephen James O'Meara. *Mars: The Lure of the Red Planet*. Amherst, N.Y.: Prometheus Books, 2001. Written for the general reader, this book explores the history of Mars observations from early telescopes through Viking and plans for future exploration as of the mid-1990's.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Mars: Illusions and Conspiracies; Mars: Possible Life; Mars's Atmosphere; Mars's Craters; Mars's Polar Caps; Mars's Satellites; Mars's Valleys; Mars's Volcanoes; Mars's Water; Meteoroids from the Moon and Mars; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits; Planetary Rotation; Planetary Satellites; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets.

Mars's Atmosphere

Categories: Mars; Planets and Planetology

Because many similarities exist among the planetary atmospheres, the study of one planet may contribute to the understanding of others. The atmosphere of Mars, with its simple structure, can be used to model certain aspects of Earth's atmosphere and is therefore a valuable aid in comprehending the past, present, and future of that atmosphere.

OVERVIEW

One goal of planetary scientists is to unravel the history of the Earth in the context of the origin and evolution of the solar system. Because all data record what is presently observed, conditions and processes of the past must be inferred from current conditions and processes. Scientists construct a model of an atmosphere's history from those inferences and use that model to project the future evolution of the atmosphere. The model is revised as better data are gathered.

Astronomers generally assume that the terrestrial planets (Mercury, Venus, Earth, and Mars) have had two atmospheres. Primitive atmospheres would have existed soon after the formation of the planets and would have been distinctly different in composition from those that exist today. These atmospheres contained hydrogen, helium, and other lightweight compounds with speeds near or above the escape velocities of terrestrial planets. Consequently, primitive atmospheres would be lost over a reasonably short time, even as the planets cooled.

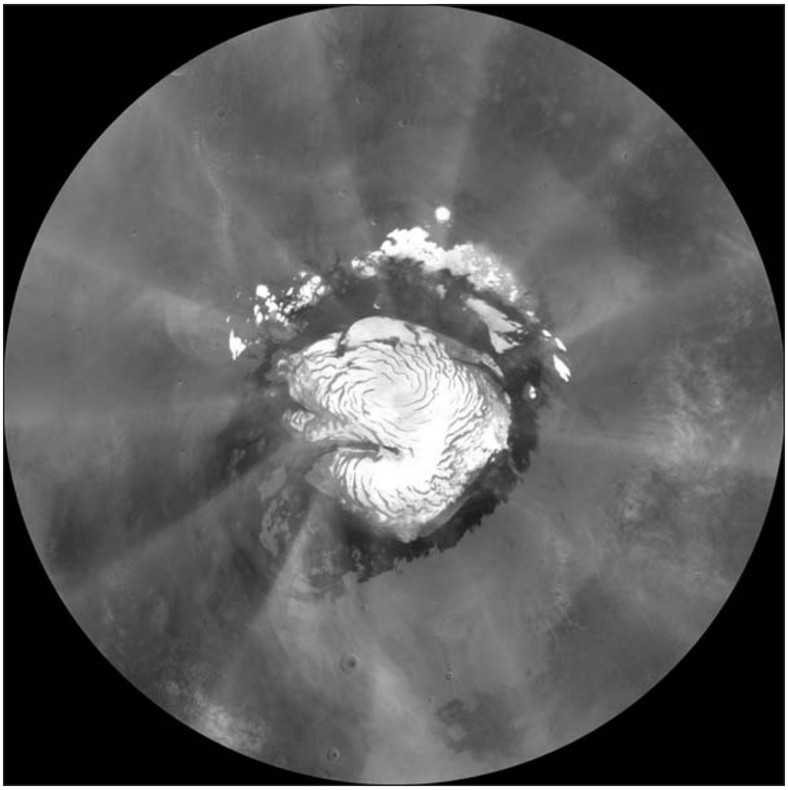
Secondary atmospheres developed as nitrogen, carbon dioxide, water, and argon were released from the planetary interiors as molten rock outgassed and volcanoes erupted. Volcanoes on Earth emit mostly carbon dioxide and water. Because scientists assume that the planets formed from the same cloud of protoplanetary material, they also assume that the gases coming from the interiors of different terrestrial planets would be similar. Therefore, the secondary atmospheres of Venus, Earth, and Mars should contain the same compounds, but the specific quantities of those compounds should vary as conditions on the planets vary. Venus should have mostly water vapor because it is nearer the Sun and has a higher surface temperature than does Earth. Mars should have mostly solid water because its distance from the Sun makes it a colder planet than Earth. Of course Earth has water in all three phases: gas, liquid, and solid.

Until the first spacecraft reached Mars, there was no accurate method for determining the pressure, composition, and temperature profile of its atmosphere. Mars was expected to have a thin atmosphere, with its low escape velocity of 5.0 kilometers per second. Most estimates of Mars's atmospheric pressure made during the first half of the twentieth century were between 1.2 and 1.8 pounds per square inch (Earth's is 14.6 pounds per square inch at sea level)—high enough for liquid water to exist on the surface of Mars as long as the temperature is not above 310 kelvins. Most scientists did not believe that Mars could retain much water, but if pressure estimates were accurate, the existence of water on Mars could not be eliminated from models of the planet and its atmosphere.

Composition of the Martian atmosphere was difficult to determine from Earth because the Earth's own atmosphere obscured much of the information that came from the planet. Because the atmospheres of the terrestrial planets were expected to have similarities, astronomers looked for nitrogen, oxygen, water, and carbon dioxide. Nitrogen is very difficult to observe, so no one was surprised when it was not detected in visible light. Astronomers, however, expected to find a significant amount of nitrogen when a spacecraft arrived at the planet. Earthbound telescopes also failed to detect oxygen or water but found carbon dioxide. Astronomers used this information and their assumptions to predict that the Martian atmosphere was largely nitrogen with some carbon dioxide present.

During the late 1960's, a series of United States Mariner spacecraft flew past Mars and found that atmospheric pressure was less than 0.09 pound per square inch. Roughly 95 percent of the atmosphere was carbon dioxide, with 1-3 percent nitrogen. Many astronomers concluded that Mars was dead and Moon-like. More recent analysis of the Martian atmosphere has shown its composition to be 95.3 percent carbon dioxide, 2.7 percent nitrogen, 0.13 percent oxygen, 0.03 percent water, and 1.6 percent argon, with trace amounts of krypton and xenon.

Average temperatures at the surface of Mars are about 215 kelvins. The warmest spot on the planet may reach 300 kelvins near local noon, but temperatures drop to less than 192 kelvins at the Viking 2 landing site. Polar regions are much colder, as confirmed in 2008 by the Mars Phoenix lander after it touched down on the Red Planet's northern polar region. Liquid water will not exist under these conditions. During the summer days, however, the surface temperature is high enough for liquid water to exist briefly. Because the vapor pressure of water in the atmosphere is very low, this liquid water evaporates quickly. As winter begins, the water molecules freeze to cold dust particles in the atmosphere. Carbon dioxide molecules also attach themselves during the cold nights, and when particles have enough mass, they fall to the surface. The temperatures during the day are high enough to vaporize carbon dioxide but not water. Pictures of the soil around the Viking 2 lander show a frost of these water-ice-coated dust grains. Clouds or ground fog of water ice crystals form about half an hour after dawn in areas heated by the Sun. Beyond latitudes of 65°, winter conditions cause carbon dioxide to freeze, and a hood of carbon dioxide clouds and haze hangs over the polar regions. The coldest regions on Mars are the poles during win-



On October 11, 2008, the Mars Reconnaissance Orbiter's Mars Color Imager took this picture of the north polar ice cap on Mars; a 37,000-kilometer-diameter dust storm is moving counterclockwise through the Mars Phoenix landing site. (NASA/JPL-Caltech/Malin Space Science Systems) ter. Polar caps made of carbon dioxide ice change size as the seasons change, but a permanent cap of water ice remains at each pole, where the temperatures never get warm enough to allow water to melt or vaporize. Mars Phoenix dug trenches into the soil near the lander and found evidence of the permafrost layer not very far beneath the surface. The permanent cap at the south pole is smaller than the one at the north pole because Mars is much nearer the Sun during the south pole's summer than during summer in the northern hemisphere. The total atmospheric pressure changes by 26 percent seasonally because of the vaporization/condensation cycle of carbon dioxide at the poles. Much of the carbon dioxide vaporized at one pole moves toward the other pole and precipitates there.

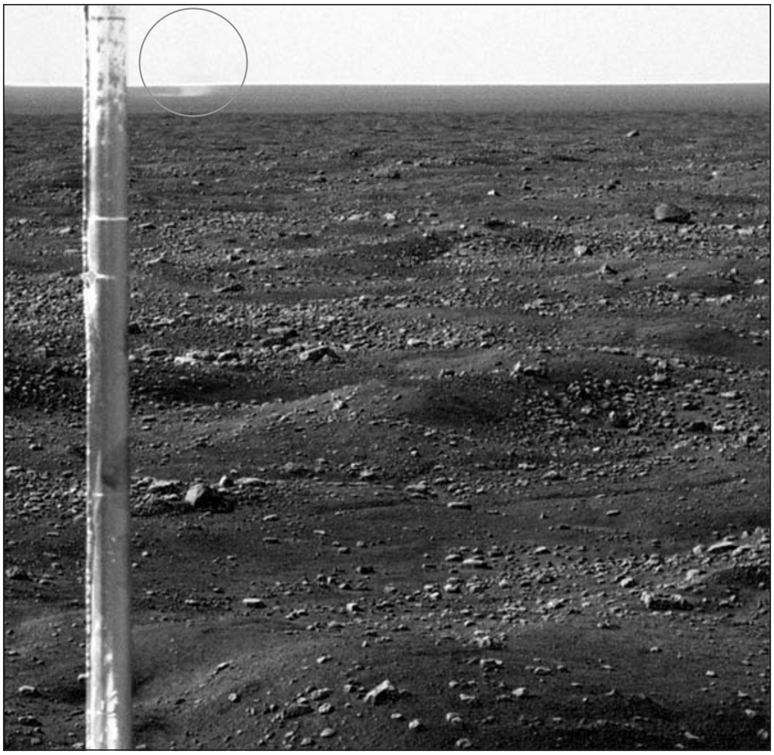
Mars's Atmosphere Compared with Earth's

	Mars	*Earth*
Surface pressure	6.1-9 millibars	1.014 bars
Surface density (kg/m^3)	~0.020	1.217
Avg. temperature (kelvins)	~215	288
Scale height (km)	11.1	8.5
Wind speeds (m/sec)	2-30	up to 100
Composition		
Argon	1.6%	9,430 ppm
Carbon dioxide	95.32%	350 ppm
Helium	—	5.24 ppm
Hydrogen	—	0.55 ppm
Hydrogen deuteride	0.85 ppm	—
Krypton	0.3 ppm	1.14 ppm
Methane	1.7 ppm	—
Neon	2.5 ppm	18.18 ppm
Nitrogen	2.7%	78.084%
Nitrogen oxide	100 ppm	—
Oxygen	0.13%	20.946
Water	210 ppm	1%
Xenon	—	0.08 ppm

Note: Composition: % = percentages; ppm = parts per million.
Source: Data are from the National Space Science Data Center, NASA/Goddard Space Flight Center.

In the northern and southern hemispheres, high-pressure systems form during summer months and low pressure develops during winter. The pressure difference is the greatest when it is summer in the southern hemisphere and winter in the northern hemisphere. Large-scale wind currents flow toward the north pole during northern winter and toward the south pole during southern winter. Dust particles picked up by these winds cause large-scale dust storms. These storms cause little erosion, because the thin Martian atmosphere can carry only small dust particles. Viking 1 measured wind gusts up to 26 meters per second as a dust storm arrived. The most vigorous storms may involve wind velocities greater than 50 meters per second. The Mars Pathfinder landing stage provided wind speed and direction data during 1997. It picked up mostly light winds, with direction varying considerably over the mission. The Mars Exploration Rovers imaged dust devils. Indeed some of the dust devils performed a useful function for the rovers as they blew off dust that had accumulated on solar panels. This cleaning effect boosted spacecraft power.

Some data imply that the Martian atmosphere was denser at one time than is presently observed. Channels (not canals) found by the Mariner spacecraft have the same appearance as channels formed on Earth by flowing water. Many craters show more erosion than is possible with the current atmosphere. Indirect evidence indicates that significant quantities of nitrogen, water, and carbon dioxide have been outgassed from the Martian interior. Many astronomers thus believe that the Martian atmosphere was once denser and warmer than now and that it became moist periodically as polar caps melted. Rivers may have flowed during these periods. As time passed, pressure gradually decreased as water vapor and carbon dioxide were lost from the atmosphere. Those losses reduced the capacity of the atmosphere to retain heat. A change in the tilt of Mars's rotational axis and a change in Mars's orbital path also reduced the temperature. As temperatures



At the Mars Phoenix Lander's site in the north polar region, a dust devil (circled) can be seen near the horizon (just right of the mast of the lander's meteorological station). (NASA/JPL-Caltech/University of Arizona/Texas A&M University)

dropped, water from the atmosphere was permanently trapped in the polar caps. Eventually, much carbon dioxide was also deposited in the polar caps, and the current cycle of vaporization and condensation was established. Exposed solid water ice was seen to sublimate away during Mars Phoenix investigations in 2008.

In mid-January, 2009, the National Aeronautics and Space Administration (NASA) announced a discovery that held the potential to become a paradigm shift in the nature of Martian exploration by future robotic spacecraft. Using ground-based telescopes at great altitude atop Hawaii's Mauna Kea, astronomers detected the infrared signature of methane in the thin Martian atmosphere. Lacking protection from solar ultraviolet radiation, methane molecules in the atmosphere would be subject to dissociation. Thus, the detection of significant amounts of methane indicated that the gas was almost certainly being replenished. Methane production could have a geological origin, but it also could have a biological one. A geological source for Martian methane could involve the conversion of iron oxide into serpentine minerals, a process that would require water, carbon dioxide, and an internal heat source; this process does occur on Earth. A biological origin for Martian methane seen in the atmosphere would involve digestion processes inside microorganisms. The methane data NASA reported could not be used to distinguish between a biological or geological source for Martian methane production. As a result, the space agency began to consider sending the next planned rover, Mars Science Laboratory, to the location of a methane vent to conduct analyses that might determine whether biology was involved in methane production on the Red Planet or if it came from a geological process.

METHODS OF STUDY

Astronomers face the challenge of collecting data on objects that are millions of miles away. Earth-based telescopes collect light that is analyzed for relevant information. This technique, however, often does not provide the precision needed for study of planets. The advent of the space program gave astronomers new opportunities to gather data as sophisticated spacecraft traveled to the remote parts of the solar system. Mariner 4 took the first close-up spacecraft photos of Mars on July 15, 1965. It was followed by Mariner 6 and 7 flybys in 1969, the Mariner 9 Orbiter in 1971, and the Viking Orbiters and Landers in 1976.

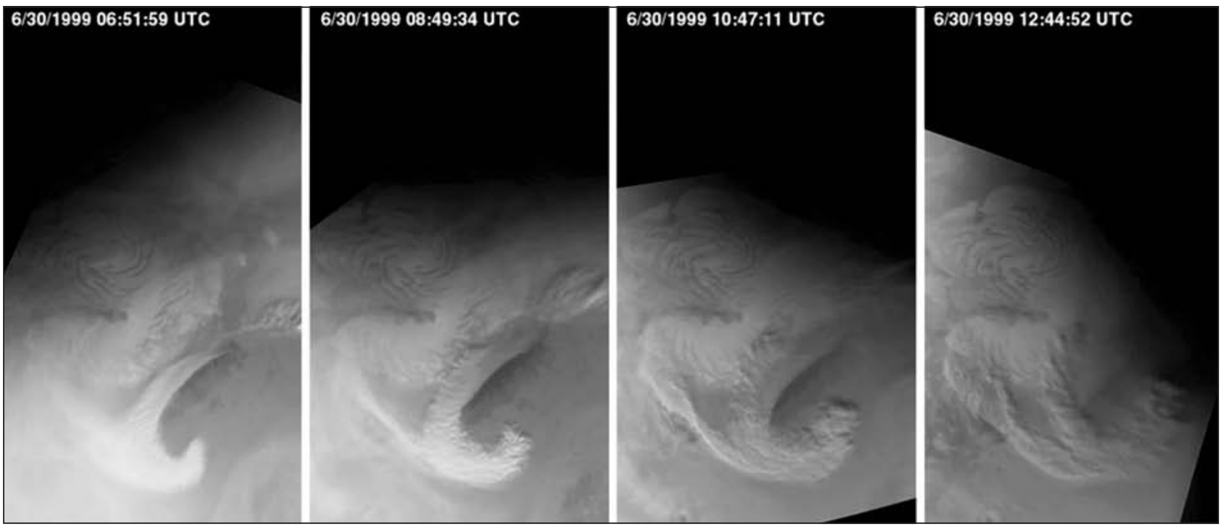
Although the Mariner spacecraft had determined that carbon dioxide was the main component of the Martian atmosphere, concentrations of other compounds present in small amounts were still unknown. As the Viking landers descended through the atmosphere, they looked for nitrogen, argon, and other elements whose molecular weight was less than 50. A mass spectrometer, an instrument specifically designed to find compounds and identify them according to their masses, analyzed the atmosphere at altitudes above 100 kilometers. Nitrogen was discovered with an abundance of roughly 3 percent. Argon, an inert gas, was found with an abundance of 1.6 percent. Another instrument, the retarding potential analyzer, showed that many hydrogen and oxygen ions were escaping from the atmosphere. Because these two elements combine to form water, their loss can be expressed in terms of water loss. Roughly 240,000 liters of water were lost each day.

The gas chromatograph mass spectrometer was used in analyzing gases emitted from experiments with the soil samples; however, it was also used to find krypton and xenon, which could not be detected by the upper atmosphere mass spectrometer because the atmospheric concentrations of these two elements were too low. Concentrations of these gases were enriched by a procedure in which Martian air was pumped into the sample chamber, from which carbon dioxide and carbon monoxide were removed. By repeating this process several times, krypton and xenon concentrations were increased to measurable levels.

The meteorology boom contained sensors to measure temperature, wind direction, and wind velocity. These sensors are thin bimetallic wires. As they change temperature, they induce an electric current. The size of this current determines the temperature. Wind cools the wires; by measuring the current passed through the wires to warm them to their original temperature, researchers can determine the velocity of the wind. Wind direction is measured by monitoring which of the sensors is cooling faster. Atmospheric pressure was measured by a sensor containing a thin metal diaphragm. The amount of movement of the diaphragm was used to determine the pressure.

CONTEXT

Planetary exploration has as its broad goals the search for life and for clues concerning the origin of the solar system. Atmospheric conditions that include the presence of water and nitrogen must exist in order for the basic materials for life to exist. The dense, hostile atmosphere of Venus destroyed any hope of finding life there. The Martian atmosphere in some ways resembles Venus's atmosphere more than Earth's; in fact, Venus and Mars each have about 95 percent carbon dioxide and 3 percent nitrogen, but atmospheric pressure on Venus is more than nine thousand times that on Mars.



The Mars Global Surveyor's cameras sent back these images of a storm over Mars's north pole in 1999 as the Martian summer was ending and autumn beginning. (NASA/JPL/Malin Space Science Systems)

Although there is evidence that water once flowed on the Martian surface and that the atmosphere was once thicker than it is now, there is no absolutely conclusive evidence that there has been enough water in the atmosphere to cause rain. Conditions for the development of life do not seem to have existed in the Martian atmosphere, but much more investigation is necessary to discount the possibility.

Further exploration of the planet will continue to be conducted by robotic spacecraft. Astronomers would also like to send scientists to explore Mars and set up a long-term research base much like those international scientists have established at the South Pole of Earth. Because the Martian atmosphere has a simple structure and because there are no bodies of water to affect the airflow, the atmospheric movement follows predictable patterns. Study of this simple system could lead to a better understanding of the more complex atmospheres of Earth and perhaps Venus. On Earth, this understanding could lead to better prediction of weather patterns, increased agricultural production, and decreased danger from natural disasters. Because carbon dioxide plays an important role in the greenhouse effect, study of the Martian atmosphere could reveal important information on how to deal with the increasing concentration of carbon dioxide in Earth's atmosphere.

The beginning of the twenty-first century saw a tremendous concern that Earth might be suffering global warming, whether from natural causes or from human impact on the planet. However, understanding of atmospheric physics was insufficient to predict with any certainty the future course of changes in Earth's climate. A more basic issue than global warming needs to be investigated before a better model of Earth's total climate system can be developed: the conditions that led to the evolution of the terrestrial planets' atmospheres. Venus, Earth, and Mars have a great deal in common. Essentially they started with similar early conditions. Why did Venus develop a thick carbon dioxide atmosphere and a runaway greenhouse effect, which left the planet's surface tremendously hot, whereas Mars became arid and cold, with a thin carbon dioxide atmosphere? Why did Earth's atmosphere develop in a way that led to the production of life as we know it?

Astronomers also believe that a better understanding of Earth's atmosphere will help them to draw conclusions about the possible existence of other planets like Earth in other solar systems. If life can form from inorganic matter, a careful study of the atmospheres on Earth and Mars could set limits on the range of conditions suitable for life to exist. Such a study may even conclude that life cannot spontaneously erupt from nonliving matter. Such a conclusion would require a total revamping of modern scientific thought.

The major question to be answered in understanding the climate and atmospheric evolution on Mars involves where its water went. If that water remains on the planet in large quantities, it could have huge impact upon future exploration of the planet.

Dennis R. Flentge

FURTHER READING

Carr, Michael H. *The Surface of Mars*. Cambridge, England: Cambridge University Press, 2007. The author provides a complete description of the geological heating of Mars as understood based on results from Mars Global Surveyor, Mars Odyssey, Mars Reconnaissance Orbiter, Mars Express, and the Mars Pathfinder and Mars Exploration Rovers. Heavily illustrated with comprehensive reference list.

Consolmagno, Guy, and Martha Schaefer. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994. Using low-level mathematics but integral calculus where required, this text demonstrates how the discipline of planetary science progresses by questioning previous understandings in light of new observations. Aimed at college students, but accessible to nonspecialists as well.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that had investigated all the planets by the publication date. Takes an astrophysical approach to give our solar system a wider context as just one member of similar systems throughout the universe.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science written from a comparative planetology perspective rather than devoting individual chapters to each planet and body in the solar system. Material on Mars covers all aspects of both Earth-based and spacecraft observations of Mars.

Kargel, Jeffrey S. *Mars: A Warmer, Wetter Planet*. New York: Springer Praxis, 2004. A member of Springer Praxis's excellent Space Exploration series. The book takes the reader on a search for Mars's water. The author provides a convincing case that the picture of a dry, waterless world portrayed initially by the early Mariner probes is not the Mars of today's understanding.

Lewis, John S. *Physics and Chemistry of the Solar System*. 2d ed. San Diego, Calif.: Academic Press, 2004. Suitable for an undergraduate course in planetary atmospheres, but accessible to the general reader with a technical background.

Moore, Patrick. *Guide to Mars*. New York: W. W. Norton, 1977. Dated but useful for its overview of the historical development of information about the planet. Describes the pre-Mariner data through the Viking missions. Some black-and-white photos; several maps. For a general audience.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/ Addison-Wesley, 2003. A discussion of data from each of the planets. Contains a large number of photographs and line drawings. Although intended as an astronomy textbook, it provides good reading for anyone with an interest in the solar system. Some sections require a science background.

Zubrin, Robert. *Entering Space: Creating a Spacefaring Civilization*. New York: Tarcher, 2000. The author displays a gung-ho attitude toward making humanity a truly spacefaring species by accepting the challenge of journeying to Mars sooner rather than later. Speculates beyond travel to Mars.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Mars: Illusions and Conspiracies; Mars: Possible Life; Mars: Surface Experiments; Mars's Craters; Mars's Polar Caps; Mars's Satellites; Mars's Valleys; Mars's Volcanoes; Mars's Water; Meteoroids from the Moon and Mars; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits; Planetary Rotation; Planetary Satellites; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets.

Mars's Craters

Categories: Mars; Planets and Planetology

Examination of Mars's craters reveals in part the history of modification of Mars's surface. Images transmitted from the first spacecraft sent to fly by Mars revealed a cratered surface somewhat resembling the lunar surface rather than one that might have been more Earth-like. However, craters on Mars undergo weathering processes that are different from those on either Earth or the Moon.

OVERVIEW

There are two classes of craters on any planet or other terrestrial body: impact craters, caused by meteors or comets crashing into the surface, and volcanic craters. On all the known bodies of the solar system, including Mars, impact craters outnumber volcanic craters by the thousands.

Early views of the planet Mercury and the detailed knowledge of Earth's moon depicted both bodies as heavily cratered. Most of this cratering occurred in the earliest 500 million years, after the planets had formed by accretion out of the primordial material that constituted the early solar nebula and Sun. During this period, the planets were heavily bombarded from space by meteors and planetesimals caught in the gravitational pull of the newly forming planets.

Unlike Mars, Mercury, and its own moon, Earth does not show evidence of having been heavily cratered, although it was subjected to the same rate of incoming meteors; Earth's thick atmosphere burns up and destroys any incoming body of less than 1,000 kilograms in mass. Earth's widespread weathering processes also quickly erase evidence of meteoritic craters. Any trace of a meteor falling into the ocean is either totally erased or largely mediated by the water. On airless, inactive bodies, however, there are few mechanisms to erase cratering, even though some occurred billions of years ago.

The biggest craters on Mars, typically caused by very large impacting bodies (such as asteroids), are called basins. The largest such basin on Mars is called Hellas Planitia; with a diameter of 1,600 kilometers and a depth of 3 kilometers, it is the deepest point on the planet's surface. (As large as it is, it is still not the largest known in the solar system. The lunar basin called Oceanus Procellarum—the Ocean of Storms—has a diameter of some 2,500 kilometers.) Other such large features are Isidis (1,400 kilometers) and Argyre (900 kilometers).

The most common type of crater found on Mars is called a rampart crater, first described by planetary geologist J. F. McCauley in 1973. Rampart cratering seems almost unique to Mars. This interesting morphology consists of a central crater wall with an ejecta pattern that resembles a resolidified flow pattern radiating outward from the crater wall as a low ridge and radial striae. This flow pattern hints that the impact actually liquefied the subsurface materials, causing them to flow away from the impact point, then resolidify as a rampart, or gently sloping wall. Causes of this fluidized ejecta are thought to have been entrapment of atmospheric gases in the ejecta or, more significantly, water in the ejecta material, such as permafrost, which acts as a lubricant, allowing the materials to flow away from the impact point. The latter concept is a favorite one of those who hope to discover immense amounts of water locked up in subsurface deposits as permafrost. Since rampart cratering is prevalent on much of Mars, such planetwide deposits of permafrost would be a very positive sign for eventual human exploration of the planet.

One of the most unusual craters found on Mars is the pedestal crater, formed by weathering action that cuts away at the base of the crater until the crater takes on the appearance of a pedestal. The unique meteorology and geology required to produce such craters are found on Mars between 40° and 60° north latitude.

In examining craters on Mars, planetary scientists are able to define the crater's age, and hence that of the surrounding terrain, through a process of observing certain of its characteristics. The process of a crater eroding away is called degradation. Examining the original impact that formed the crater is a starting point; a new crater with little or no degradation is one with sharp edges and fresh ejecta outlying its central diameter. As a crater degrades through natural processes, its rim loses sharpness, and crater walls slump and lose their definition. Ultimately, the walls may form gullies and eventually become completely degraded to surface level. In the final stage, the crater is hardly noticeable over the terrain, and some planetary scientists even call such a crater a ghost.

On Earth, the most profound degradation process is weathering, which can erase even large craters relatively quickly, in geologic terms. On Mars, such processes occur much more slowly. Since liquid water does not presently flow on the Martian surface, weathering processes are confined to wind erosion, or aeolian (wind) processes. It is estimated that aeolian erosion is responsible for filling in craters on Mars at the rate of 0.0001 centimeter per year. From these crater studies, there is considerable evidence that such erosion has slowed in the planet's recent geologic past. The last period of relatively heavier erosion occurred roughly 600 million years ago, according to some crater studies.

The interior morphology of Martian craters is highly variable. It is typically dependent on the mass of the incoming projectile and the composition of the impact site. In smaller craters, the bottom usually assumes a more spherical shape. In larger craters, the central region of the crater flattens out until, in craters greater than 25 kilometers in diameter, a central peak is formed. Such flatness is often caused by the impacting body liquefying the crust or by

The High Resolution Imaging Science Experiment (HiRISE) relayed this image of the Mars Phoenix Lander, attached to its parachute (enlarged view lower left corner), as it drifted down to the landing site in Mars's arctic region on May 25, 2008. The 10-kilometer-diameter Heimdall Crater occupies most of the image area. (NASA/JPL-Caltech/University of Arizona)

magma welling up from the planet's interior to fill in the crater.

By examining craters formed within other craters in a process called superimposition, planetary geologists are able to determine the age of whole regions of Mars's surface. With an overall planetary comparison, which includes mass crater counts, it has been discovered that there are areas that have changed little since Mars was formed, while other regions, notably volcanic regions, have changed in the recent geologic past.

In some of the most fascinating crater studies accomplished to date, scientists have attempted to determine the age of what appears to be massive river channels on Mars. Although conditions on Mars at present will not allow liquid water to exist on the planet's surface, it appears that water once flowed in an extensive series of channels. Planetary geologists have examined craters that overlie the enigmatic channels. They range in age from 3.5 billion years to as young as fewer than 200 million years, strongly suggesting that the channels have been formed on a cyclic basis throughout the geologic history of Mars.

Mars has the largest volcanic craters in the known solar system, resulting from volcanoes between ten and one hundred times larger than the largest volcano on Earth, Mauna Loa in Hawaii. Four very large volcanic craters exist in close proximity in the Tharsis region. The largest volcano is Olympus Mons, a 200-million-year-old volcano with a crater 80 kilometers across. These Martian volcanoes, called shield volcanoes after their earthly counterparts, have craters nested inside one another. Three other Tharsis volcanoes are roughly alike in size, all much larger than Mauna Loa. The crater on Arsia Mons is 140 kilometers in diameter (compared to Mauna Loa's 2.8 kilometers. The largest Pavonis Mons crater is 45 kilometers in diameter and 5 kilometers deep (compared to Mauna Loa's 0.2 kilometer of depth). The largest of the Ascraeus Mons craters is roughly the same diameter as the largest Pavonis Mons crater. The volcanic plains spreading away from the Martian volcanoes are dated according to crater distribution. The youngest is Pavonis Mons, at 80 million years.

METHODS OF STUDY

Earth-based telescopic observations of Mars cannot clearly reveal its distinct surface features. The distance is too great and the combined effects of Earth's and Mars's atmospheres reveal only indistinct and blurred splotches of color. The means for collecting aggregate crater data is basically orbital photography from spacecraft. Almost all such data have come from flyby spacecraft, orbiters, and landers.

The first distinct images of the planet's surface came from NASA's Mariner 4 spacecraft on July 14, 1965. This robotic spacecraft transmitted nineteen photographs of the Martian surface back to Earth. In the twenty-two minutes that it took to transmit the photographs, previous speculation about the Martian surface was laid to rest. The most surprising and pronounced of the Mariner 4 revelations was the extensive cratering of the Martian surface, a finding few planetologists had expected. Popular interest in Mars waned at this point, however, because the planet appeared to be rather Moon-like instead of perhaps a place where life might have had a chance to develop.

Based on observations from Earth-based telescopes, it was widely expected before 1965 that there was enough weathering on Mars to have erased most of its primordial cratering. Mariner 4 clearly showed that this was not the case. The most immediate implication was that Mars was far less active meteorologically and geologically than was initially thought. Later, both extreme views were mediated by data from spacecraft that gave rise to more detailed studies, which showed that Mars has an active weathering process, though not as vigorous as that of Earth.

The twin Mariner 6 and 7 spacecraft flew past the Red Planet only a few weeks after the Apollo 11 astronauts returned to Earth following the first crewed landing on the Moon. At this point, interest in sending humans to Mars was at a precarious stage: Some were thrilled that Apollo had achieved its goal but now were no longer interested in pursuing expanded human-based scientific exploration of space; others were eager to build on the success of Apollo and fly to Mars during the 1980's. The supposition of many was that the picture of Mars as revealed by Mariner 4 was atypical of the entire planet. However, both Mariners 6 and 7 sent back a catalog of images of the Red Planet that dashed the hopes of those expecting Mars to be more Earth-like. The majority of the images displayed cratered features with no evidence of surface water.

NASA program designers in the 1960's included sending a pair of spacecraft toward a planetary target to ensure mission success if a launch or other accident resulted in the loss of one. The wisdom of that policy had been demonstrated when Mariner 1 suffered a launch accident but Mariner 2 encountered Venus. In early Mars exploration, Mariner 3 had failed but Mariner 4 was a huge success. Both Mariners 6 and 7 were successful, but of the pair of orbiters in the Mariner Mars next sent to Mars by NASA, Mariner 8 suffered an upper stage shutdown 265 seconds after liftoff and fell back into the Atlantic Ocean. Mariner 9, however, became the first spacecraft to enter orbit about another planet on November 13, 1971.

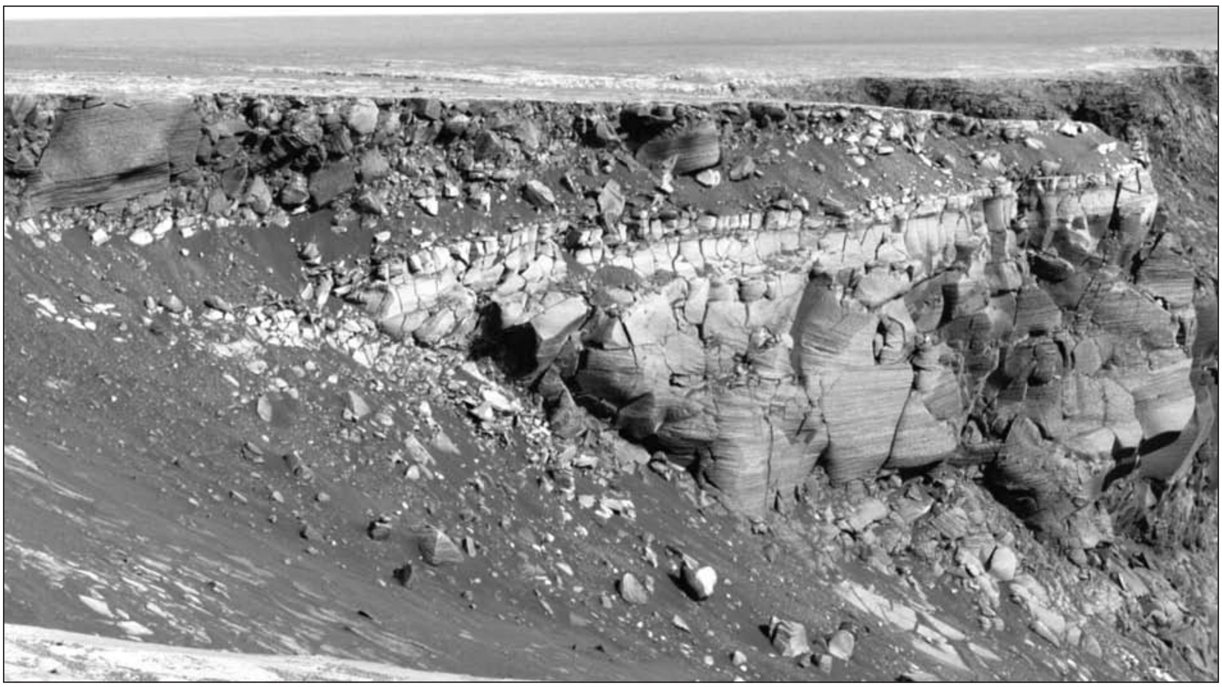
Whereas Mariners 4, 6, and 7 had revealed an inactive world more like the Moon, pockmarked with craters, Mariner 9 returned images with resolutions as good as 0.1 kilometer per pixel and thus revealed an entirely new side of Mars's surface characteristics. In addition to craters, Mariner 9 provided evidence of large-scale flows of water across the planet's surface in the distant past. Volcanic craters were imaged to greater resolution, but no evidence was obtained to indicate present-day volcanic activity or hot spots near Mars's huge volcanoes. The images and data from Mariner 9 helped scientists determine how and when to attempt the next stage of Mars exploration: sending landers to the surface in 1976.

Viking 1 and 2 each consisted of a joint orbiter and lander. Whereas the orbiters continued to image as much of the Martian surface as possible—including craters, valleys, volcanoes, and polar regions—the primary focus of the landers was to search for conditions that might support in the present, or might have supported in the past, primitive life on Mars. Viking 1 landed on Mars on July 20, 1976, in western Chryse Planitia, and Viking 2 on September 3, approximately 200 kilometers to the west of the crater Mie in Utopia Planitia. Both landers analyzed the soil but failed to find any organic material or evidence of life. It would be twenty-two years before another successful powered landing of a spacecraft on the Martian soil. Mars Phoenix successfully touched down on the Red Planet's northern arctic region on May 25, 2008, in an area largely devoid of large rocks. The landing site was chosen because it offered a high possibility of the spacecraft's encountering permafrost with a layer of ice either on the surface or not very deep beneath.

Russian plans for extensive exploration of Mars ended with the Phobos 1 and 2 spacecraft in 1988 and 1989. One failed en route to Mars, the other as it approached the satellite Phobos. With the dissolution of the Soviet Union in 1991, Russian plans for Mars programs dissolved as well.

Between the Viking (1976) and Mars Phoenix (2008) landings, NASA refined its plans for Mars exploration at least twice. As a result of the indication that water had once flowed across the cratered Martian surface, a new wave of Mars spacecraft were sent to the Red Planet. It began with Mars Observer, which unfortunately exploded as it prepared to execute its Mars orbit insertion burn. Mars Observer was followed by Mars Pathfinder, a surface rover; Mars Global Surveyor, an orbiter; Mars Climate Orbiter, which failed to achieve orbit and burned up in Mars's thin atmosphere; Mars Polar Lander, which crashed in the southern polar region; Mars Odyssey, an orbiter; twin Mars Exploration Rovers named Spirit and Opportunity; the European Space Agency's Mars Express; the Mars Reconnaissance Orbiter; and the Mars Phoenix, a powered surface lander. The thrust of scientific investigation of all of these second-generation Mars probes was a search for Martian water, but that required investigation of Martian geology. This search included the possibility of water ice deposits in craters and the erosional aspect of water on crater walls.

In the process, these spacecraft provided greatly improved images of Mars. Some re-



Victoria Crater, about 800 meters in diameter, near the Martian equator, as imaged by the Mars Exploration Rover on May 6, 2007. (NASA/JPL/Cornell University) turned data over long periods of time, including detailed meteorological observations from both orbit and the surface. From this information, Martian cratering and morphology were assessed in detail. Some of the most elemental assessments of crater morphology include crater size, composition of ejecta, composition and behavior of impacted terrain, modification processes of the crater, and its effect on surrounding craters and terrain. These findings led to a determination of the ages of the craters, surrounding crater fields, and even the impacted terrain. Assessments were made of crater sizes, numbers, distribution, and ages within selected planetary areas.

Techniques for analysis of Martian craters vary with the aim of the study. For example, a statistical counting method borrowed from other applications, called a frequency distribution, is used to determine the number of craters located over a given area. This information may lead to determination of the age of the area under study, given a uniform crater deposition rate. Other statistical counting methods include cumulative distribution, logarithmic incremental distributions, and incremental frequency distributions, all of which are specific methods of presenting collected crater population data. Some appear as graphs, others as numerical tables, and others as simple maps showing locations of craters over a given area. In the largest distribution sampling exhibit, the entire planet of Mars is presented with its crater distributions marked. All of these are unique representations designed to give the researcher information in a specific way so that the concept under study, such as the age of an area, can be efficiently gauged. Inferences are made from these statistical presentations about such broad concepts as crater production and erosion.

Planetologists have arrived at a rate of cratering for Mars that is equivalent to three hundred meteors, of one kilogram or larger, striking each square kilometer every million years. From this baseline of impacts, current photographs can be compared to establish which areas are very old and which are younger geological formations. The logical extension of such knowledge is the ability to date such formations as plains, volcanic flows, and streambeds, as well as to establish a baseline for regional aeolian erosion.

CONTEXT

The study of Martian craters allows planetary scientists to determine from orbital photographs many varied characteristics of Mars. From information relating present conditions on Mars all the way back to its earliest geological history, crater studies have allowed much of the planet's chronology to be traced. Orbital studies of Mars's craters may lead to answers about the planet's most significant geological history, such as whether water ever existed on Mars in liquid form, what formed the vast river channels and canyons, and what happened to the water. Answers to such questions will determine what humanity will have to do to survive on Mars should this neighbor planet be visited and perhaps colonized, as proposed in the Bush administration's 2004 Vision for Space Exploration.

Crater studies also address questions of importance relative to Earth's own history and future. Planetologists seek to understand why Mars is so different from Earth and what indeed happened to its once apparently abundant water. Such questions of planetary history relate to what may one day happen on Earth. If the scientific community is eventually able to draw enough parallels from the study of other planets, it may be able to apply rigid mathematical projections of Earth's own future based on present conditions and trends.

Dennis Chamberland

FURTHER READING

American Geophysical Union. *Scientific Results of the Viking Project*. Washington, D.C.: Author, 1978. A compendium of articles about the Viking project, originally published in the *Journal of Geophysical Research*. It consists of detailed assessments of data from the Viking landers and orbiters. Suitable for college-level readers.

Barlow, Nadine. *Mars: An Introduction to Its Interior, Surface, and Atmosphere*. Cambridge, England: Cambridge University Press, 2008. An interdisciplinary text including contemporary data from the Mars Exploration Rovers and Mars Express. Each chapter contains necessary background information. A good reference for planetary science students.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. Filled with color diagrams and photographs, this popular work on solar-system astronomy covers planetary exploration through the Mars Pathfinder and Galileo missions. Accessible to the astronomy enthusiast.

Collins, Michael. *Mission to Mars*. New York: Grove Weidenfeld, 1990. Apollo 11 astronaut Collins provides an astronaut's vision of a trip to Mars. Examines the problems to be overcome to make such a journey possible with space-shuttle-era technology.

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001. A challenging and thorough text for students of planetary geology. Covers extrasolar planets as well as an in-depth explanation of solar-system formation and evolution. An excellent reference for the most serious reader with a strong science background.

Ezell, Edward, and Linda Ezell. *On Mars: Explorations of the Red Planet, 1958-1978*. NASA SP-4212. Washington, D.C.: Government Printing Office, 1984. The classic official NASA history of the Viking program from the original ideas in 1958 to the culmination of the project some twenty years later. Provides a detailed assessment of the political history and discusses details of the instruments that photographed Mars from orbit as well as the surface instruments that measured the winds. Generally nontechnical and accessible to most readers.

Harland, David M. *Water and the Search for Life on Mars*. New York: Springer Praxis, 2005. A historical review of telescopic and spacecraft observations of the Red Planet up through the Spirit and Opportunity rovers. Covers all aspects of Mars exploration, with emphasis on the search for water.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science. The material on Mars covers all aspects of ground-based and spacecraft observations of Mars. Takes a comparative planetology approach instead of including separate chapters on each major solar-system object.

Squyres, Steve. *Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet*. New York: Hyperion, 2006. Written by the principal investigator for the Mars Exploration Rovers Spirit and Opportunity, this fascinating book provides a general audience with a behind-the-scenes look at how robotic missions to the planets are planned, funded, developed, and flown. A personal story of excitement, frustrations, a scientist's life during a mission, the satisfaction of overcoming difficulties, and the ongoing thrills of discovery.

Zubrin, Robert. *Entering Space: Creating a Spacefaring Civilization*. New York: Tarcher, 2000. The author displays a gung-ho attitude toward making humanity a truly spacefaring species by accepting the challenge of journeying to Mars sooner rather than later using contemporary technology and daring innovation. Speculates beyond travel to Mars.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Mars: Illusions and Conspiracies; Mars: Possible Life; Mars: Surface Experiments; Mars's Atmosphere; Mars's Polar Caps; Mars's Satellites; Mars's Valleys; Mars's Volcanoes; Mars's Water; Meteoroids from the Moon and Mars; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits; Planetary Rotation; Planetary Satellites; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets.

Mars's Polar Caps

Categories: Mars; Planets and Planetology

In 1666, Gian Domenico Cassini observed Mars's surface and described two polar caps, the planet's most visible features. These caps have been studied extensively since, both telescopically and by means of the Mariner, Viking, Mars Global Surveyor, Mars Odyssey, Mars Reconnaissance Orbiter, and Mars Express space probes.

OVERVIEW

When Gian Domenico Cassini observed the polar caps of Mars in 1666, little was known about the planet's surface features. A decade before Cassini, the Dutch astronomer Christiaan Huygens suggested the presence of polar caps on Mars, but it was not until 1672 that he actually saw the planet's south polar cap, Mars's most apparent feature when one views it through a telescope. No one before Cassini had seen the polar caps in detail, because no instrument had existed to allow sufficient magnification.

After the invention of the telescope, early astronomers who observed some of the surface features of Mars reached conclusions based on analogies with Earth, the only planet whose atmosphere and chemical composition they knew in detail. It is not surprising, then, that the Martian polar caps were presumed to be like the polar caps found on Earth: composed of ice, snow, hoarfrost, or a combination of these forms of water. This theory presupposes that Mars has water, considered a requisite for life, and therefore the individual components of water, hydrogen and oxygen. This notion gave rise to the popular theory that Mars could support life in some form, a theory that was subsequently brought into considerable question and that remained controversial after the early Mariner and Viking missions to Mars.

In 1719, the Italian astronomer Giacomo Maraldi discovered that the polar caps are not centered at Mars's precise geographical poles. Maraldi also divined from his observations that the south polar region of Mars has a much greater mass than its northern counterpart. Subsequent research and observation have substantiated Maraldi's theories, revealing that the planet's north polar cap lies about 64 kilometers from its geographical north pole and that the south polar cap lies some 400 kilometers from the geographical south pole. This phenomenon occurs because of the planet's elongated orbit.

Italian astronomer Giovanni Schiaparelli trained telescopes on Mars repeatedly during the 1877 opposition of the Red Planet and drew highly detailed maps of Mars's surface. Those maps documented networks of linear features Schiaparelli was convinced existed in great number between 60° north and south in latitude. Reports of these linear features were taken out of context in public accounts. Schiaparelli had termed these features *canali*, which from Italian translates as "channels." However, these linear features were most often referred to as "canals," with the obvious implication that they were artificial structures (in turn implying that they were built by intelligent life-forms) to convey water from one location on the planet to another. Thus the notion of canals on Mars was born and proceeded to take on a life of its own in science fiction as well as in pseudoscience.

The wealthy astronomer Percival Lowell developed a fascination with the planet Mars. He built the Lowell Observatory in Flagstaff, Arizona, atop Mars Hill and spent the final twenty-three years of his life deeply engaged in astronomical investigations, especially focusing on Mars. Lowell sought to use advanced telescope technologies to search for the canals of Mars prominent in Schiaparelli's maps of the Red Planet. Lowell believed in and popularized with the public the notion that the canals were evidence of alien technology. In three books, Lowell advanced the theory that the Martians had built an elaborate system of canals to bring precious water from the polar ice caps to population centers suffering from increasingly arid conditions.

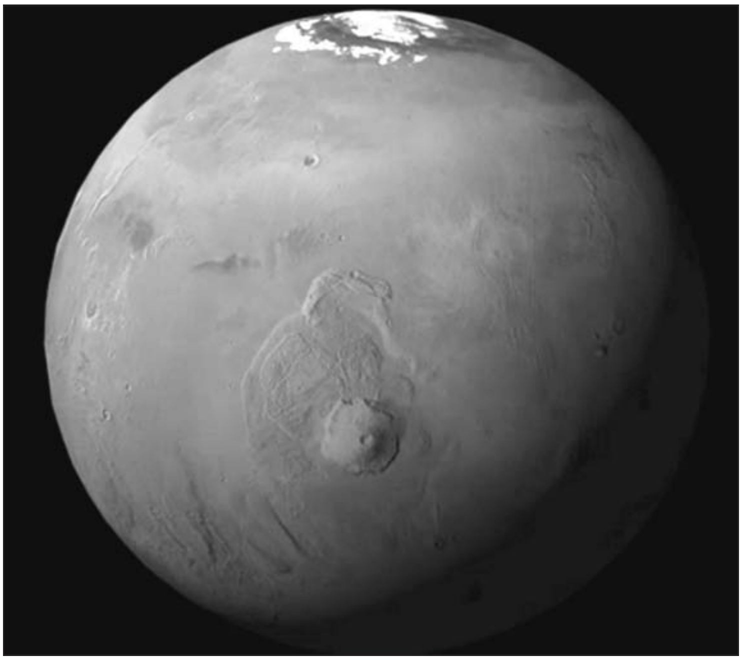
Lowell's notions were not well received by the scientific community (although he later would expend considerable effort to locate a body beyond Neptune, which bore fruit after his death with the discovery of Pluto by Clyde Tombaugh). The present atmosphere of Mars is known to be such that there can be no accumulated areas of free liquid water, such as lakes, ponds, rivers, or oceans on the planet's surface. Much evidence suggests, however, that in the past the Martian surface was extensively cut by flowing liquid. Most evidence demolishes the theory that this liquid was lava flowing from the planet's once active volcanoes. Current theory indicates, rather, that in one stage of the planet's evolution, water in liquid form was plentiful.

When everything astronomers knew about Mars came from their observations through telescopes, they could gather substantial information about the polar caps; they were at odds, however, in their interpretations of these data. They could not always be sure what they were seeing. They also had little means of knowing with any certainty the depth of the polar caps. Some scientists thought that they were hundreds of feet deep; others thought the caps were merely thin coverings of hoarfrost. Common sense supported the latter idea. The argument was that if a thin layer of hoarfrost covered the poles, it would, under warmer conditions, condense and form clouds. On the other hand, if the polar caps were composed of fairly deep layers of solid water, where would the melting residue run during the warmer season? The atmosphere of Mars was then thought to be almost totally dry, yet observations through telescopes clearly showed that in the summer, the polar caps seemed to melt partially. Certainly they diminished in size, and the areas surrounding them assumed a dark coloration, suggesting that water was mixing with minerals and dust at the periphery of the thawing area and causing what looked like a moist condition.

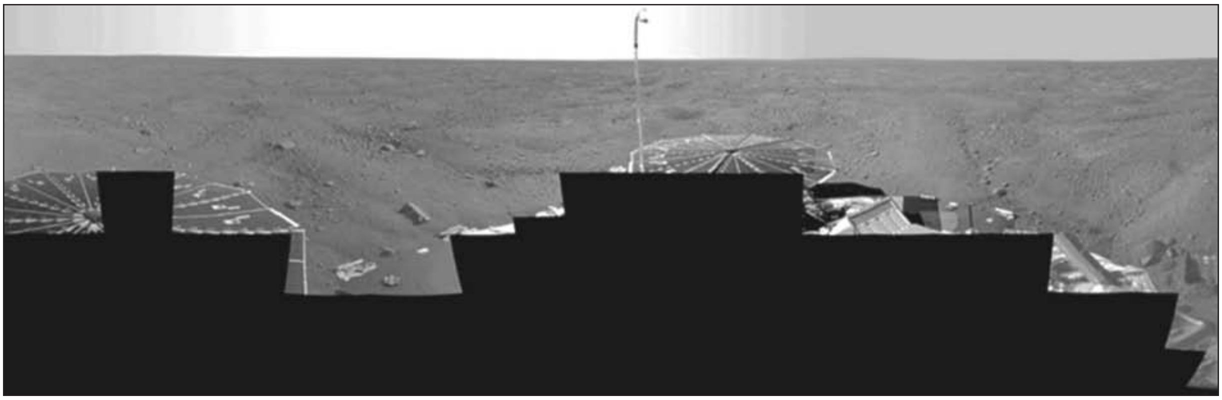
The south polar cap covers an area of more than 10 million square kilometers in winter and at times extends almost halfway to the Martian equator. Even a minimal thawing would produce water in great quantity, especially if the water ice ran to great depths; yet, no water from the polar cap appeared through the telescope to have penetrated other areas of the planet, which seemed dusty and dry.

Faced with this anomaly, scientists could do little to verify their theories until a different sort of evidence, the kind first provided by the Mariner and Viking missions, became available in quantity between 1965 and 1977. Viking 2 measured some of the temperatures of the north polar cap with a high degree of certainty and discovered that temperatures in its dark surrounding areas had a range of between 235 and 240 kelvins, and that the white, presumably frozen areas had a range between 210 and 205 kelvins. These temperatures exceed 194 kelvins, the point at which carbon dioxide sublimates into a solid.

The atmosphere of Mars, thought until the Mariner 4 flyby in 1965 to be about 85 millibars, was proved to be a thin 10 millibars; Earth's at-



Olympus Mons, the tallest volcano on Mars, appears as a bump in the lower center of this image from the Mars Global Surveyor, while the north polar region displays an icy crown. (NASA/JPL-Caltech/University of Arizona)



This panorama of a plain in Mars's north polar region shows a 360° view around the Mars Phoenix Lander and surface patterns reminiscent of those seen in permafrost areas on Earth. (NASA/JPL-Caltech/University of Arizona/Texas A&M University)

mosphere at sea level is 1,000 millibars, or 1 bar. The Martian atmosphere is comparable to Earth's atmosphere at about 24,000 meters, where the pull of gravity is all but lost. Such an atmosphere does not permit freestanding water, indicating that if the Martian atmosphere had always been the way it is now, the planet could never have had water, and the polar caps would necessarily be composed of some other substance.

This sort of thinking gave rise to the theory that the polar caps were composed of solid (frozen) carbon dioxide, or dry ice. Lowell had advanced such a theory as early as 1895; other eminent astronomers considered it more likely than the water theory as late as 1971, when Mariner 9 was uncovering data that would soon vitiate, although not completely eliminate, the carbon dioxide theory. Mariner 4, when it made its flyby of Mars in 1965, had returned data suggesting that the Martian atmosphere was composed largely of carbon dioxide under weak pressure. This information caused some astronomers to cast their lots with Lowell's theory that carbon dioxide in its frozen state covered the polar caps but that, when the temperature rose, it became, through sublimation, gaseous and returned to the atmosphere as mist or fog. This explanation helped to account for the haze often observed over the polar regions.

Mariner 9 data substantiated the theory that Mars's present atmosphere accommodates no accumulations of liquid water. It also presented incontrovertible evidence that the polar caps are composed largely of ice and are minimally 0.8 kilometer deep, indicating clearly that the Martian climate has changed through the eons and that it was once such that liquid water, now locked in the polar caps in solid form, was abundant.

The earliest space missions carried out research only concerning the north polar cap, but they managed to dispel a substantial number of misconceptions about Mars, among them the mistaken idea that the planet is totally dry. Data sent back suggested that in the northern latitudes above 60°, the atmosphere is quite moist; the atmosphere over the north polar cap has twenty times the water vapor that the atmosphere over the equatorial regions contains. During the Martian summer, surface ice exposed to the Sun evaporates in the morning, causing a mist that seems to condense, resulting later in the day in precipitation. Although these data were gathered exclusively over the north polar cap, there was little reason to think that the south polar cap differed significantly. At perihelion, the planet's closest approach to the Sun, the north polar cap does not face the Sun directly but is tilted away from it. This tilt prevents it from getting the full impact of the Sun's rays, which would, presumably, melt a polar cap composed of frozen water.

Space probes to Mars indicated that in the ancient past, water must have been quite plentiful on the planet. The climate of Mars must be viewed over eons; the planet is now much drier than it once was, and scientists think that it could evolve through this period to one in which conditions resemble what they were when the Martian atmosphere allowed the accumulation of water in bodies similar to the ones on Earth.

It is now widely accepted that portions of both of the major early theories about the polar caps were true. Although the evidence is strong that deep layers of water ice cover the polar regions and although it is known that the deep crater Korolev in the north polar cap is filled with water ice, many astronomers think that during certain seasons there are thin coatings of carbon dioxide that sublimate into the gaseous state and cause the clouds or mists that have been observed over the polar caps.

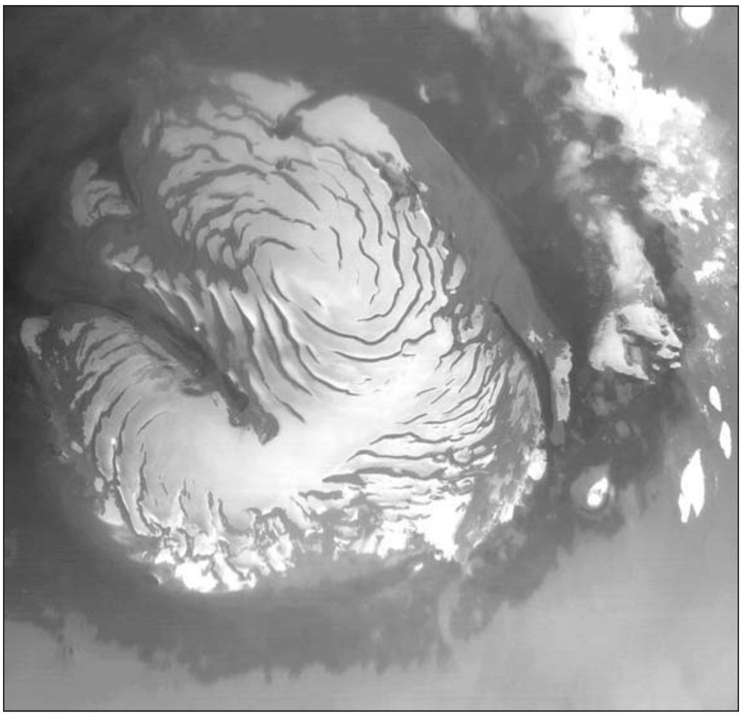
METHODS OF STUDY

In the three and a half centuries that the polar caps of Mars have been observed, descriptions of the caps have moved from highly speculative to soundly scientific. In this period, astronomers have moved from primitive optical telescopes to incredibly complex telescopes of enormous size and capable of detecting invisible radiation, strategically placed to focus on the sky and on the planets. Mars has been the most intriguing planet for most astronomers to explore, because it is sometimes a relatively close 56 million kilometers from Earth. It is also the planet that most resembles Earth in its surface features, although its atmosphere currently precludes advanced life.

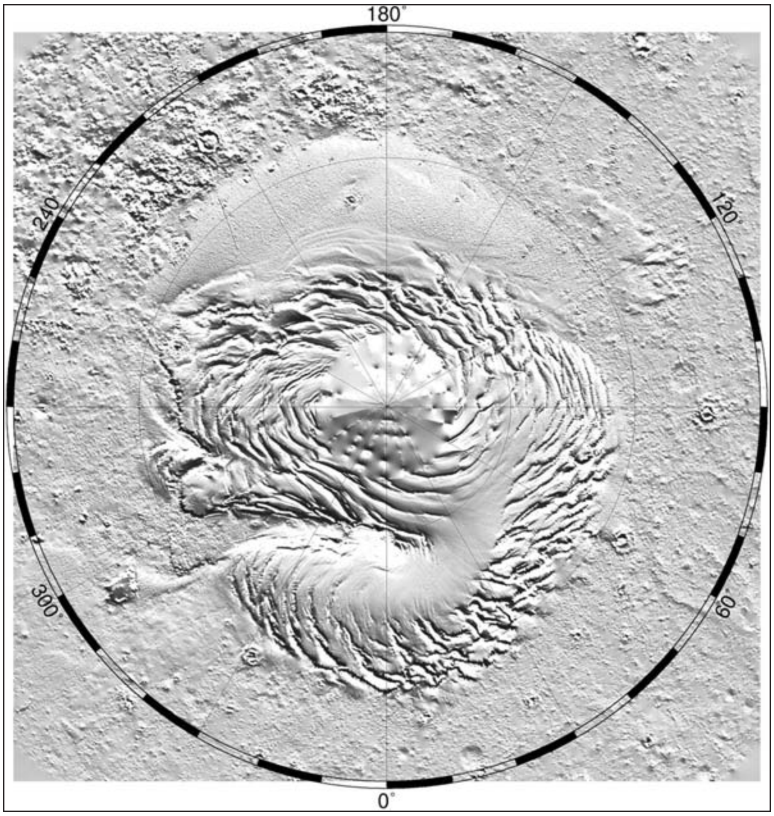
The Mariner and Viking missions first solidified human knowledge of Mars. Telescopic evidence suggested to many that, because of reflections detected from the polar caps, the caps must be composed of water ice. When the earlier Mariner missions presented evidence favoring the carbon dioxide theory, astronomers were forced to rethink their earlier stands. Later expeditions, however, offered convincing evidence in favor of the water ice theory. When Viking 1 landed, it not only photographed the surface extensively and transmitted the pictures to Earth but also deployed an arm that dug into the Martian surface and analyzed the composition of the materials it uncovered. Finding water locked in Martian rocks established clearly the former existence of water on the planet.

Mariner 9 was placed into orbit around Mars in 1971, and it sent back more than seven thousand pictures. It photographed the south polar cap continuously from November, 1971, until March, 1972, capturing the waning of the polar cap as summer advanced. These pictures provided extremely varied information and reiterated the water ice theory.

The goal of the Viking program was to land on the surface of the planet and analyze the soil for signs of life. However, it is often overlooked that the landers were each transported to Mars by an orbiter. The Viking 1 orbiter lasted from orbital insertion on June 19, 1976, until August 17, 1980, when it ran out of attitude control propellant. The Viking 2 orbiter lasted from August 7, 1976, until July 25, 1978, at which point it



Mars's North Pole, as imaged by the Mars Global Surveyor. (NASA)



This model displays the topography of Mars's north polar region and ice cap. (NASA/JPL/GSFC)

suffered a fuel leak. During those periods, these orbiters produced extensive catalogs of images of a large majority of the Martian surface, including photographs of both polar regions, which over time displayed seasonal changes.

The Mars Observer spacecraft was launched on September 25, 1992, and was designed, among other things, to generate a global distribution map of Mars's elements and minerals. This spacecraft was one of the largest and most sophisticated (for its time) to be dispatched to Mars. Unfortunately, before it could enter orbit and perform any scientific operations, including studies of the polar caps, Mars Observer was lost on August 21, 1993. However, its loss resulted in a new approach to interplanetary exploration, National Aeronautics and Space Administration (NASA) Administrator Daniel S. Goldin's infamous "faster, better, cheaper" policy, which sought to replace expensive, large-scale probes such as the Mars Observer with more focused, less expensive probes sent out into the solar system with greater frequency and more risk by using cutting-edge technologies. "Faster, better, cheaper" turned out, as evidenced by the subsequent losses of Mars Climate Orbiter and Mars Polar Lander and others to be neither better, faster, nor even cheaper in the long run. Goldin's approach managed to delay by almost a decade what had been a well-thought-out program of Mars exploration intended to robotically return rock and soil samples to Earth by the middle of the first decade of the twenty-first century. The impact of Mars Observer also had serious repercussions for studies of the Martian polar regions.

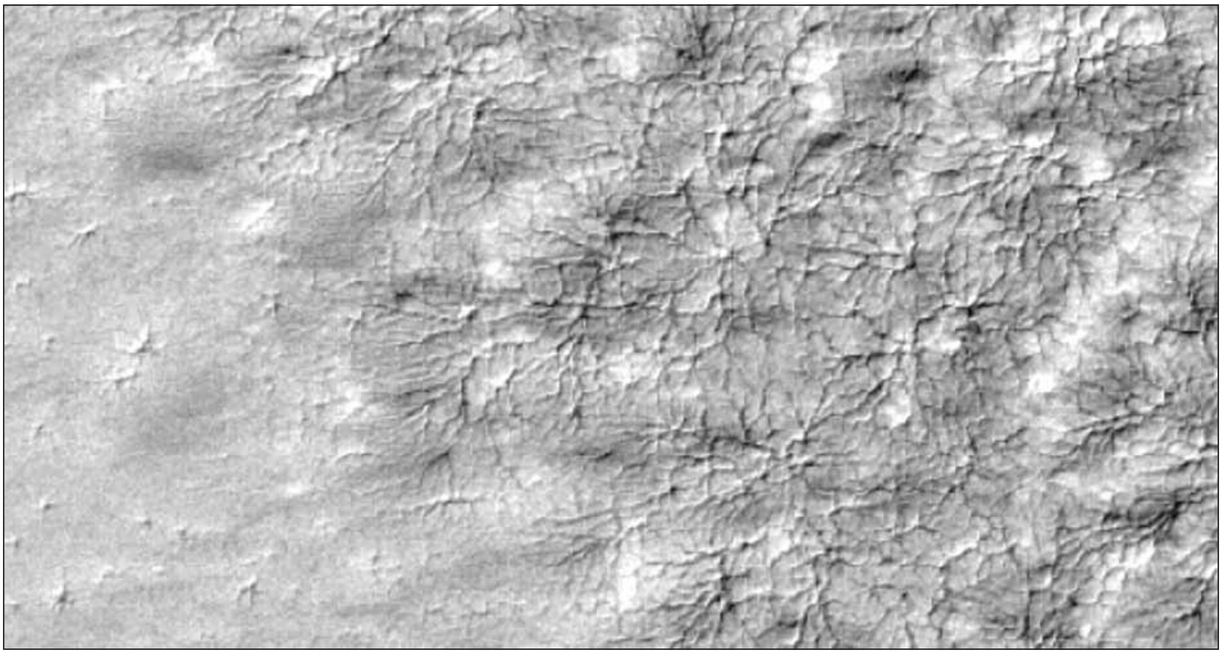
Mars Global Surveyor (MGS) was launched on November 7, 1996, and entered a stable Martian orbit on September 11, 1997. It was the first successful American spacecraft to the Red Planet in twenty years since the two Vikings. MGS entered a near-polar orbit, so that as it imaged geological features across the planet it also was used to study the polar ice caps. Its purpose, to search for evidence of water, went far beyond merely searching for water in the polar regions. MGS provided images of gullies on the walls of craters that could be the result of groundwater eroding the crater walls before freezing. In early November, 2006, MGS went silent after experiencing a problem in directing its solar arrays toward the Sun to generate sufficient electrical power.

The Mars Climate Orbiter (MCO) was NASA's next probe sent to the Red Planet after the highly successful Mars Pathfinder rover, which generated tremendous national and international excitement in 1997. MCO was launched on December 11, 1998, to study Mars's atmosphere primarily, but also, by investigating the transport of water and atmospheric circulation, it would have indirectly increased humanity's understanding of the polar caps. Unfortunately, the result of a human miscommunication between engineers, some working using the metric system and others using British units, led to MCO's coming in too close to Mars on final approach to orbital insertion. On September, 23, 1999, MCO performed its Mars orbit insertion burn but never came around the back side of the Red Planet. Entering orbit at far too low an altitude, only 57 kilometers above the surface, MCO burned up and was lost.

The Mars Polar Lander (MPL) was a powered lander intended to touch down on the surface of the Red Planet in the southern polar region in order to search for evidence of water in the soil. The spacecraft was equipped with a sample scoop at the end of a robotic arm to dig up samples. A gas analyzer was included among its scientific instruments in order to heat soil samples and detect any volatile gases, such as water vapor, given off in the process. On December 3, 1999, MPL began its entry phase at a speed of 68 kilometers per second. Unfortunately, the onboard guidance system appeared to cut off the engines, slowing its speed at an altitude of 40 meters and causing the lander to crash on the Martian surface.

Mars Odyssey was launched on April 7, 2001, and entered a nearly polar orbit on October 24, 2001. In addition to being a scientific research platform, the spacecraft was intended to serve as a communications relay for later surface probes (such as Mars Phoenix). Designed primarily to search for evidence of water, perhaps the most important discovery made by Mars Odyssey was detection of enormous amounts of hydrogen, which was interpreted as an indication of subsurface water. An extended mission was approved to continue observations of the polar ice caps for contemporary changes.

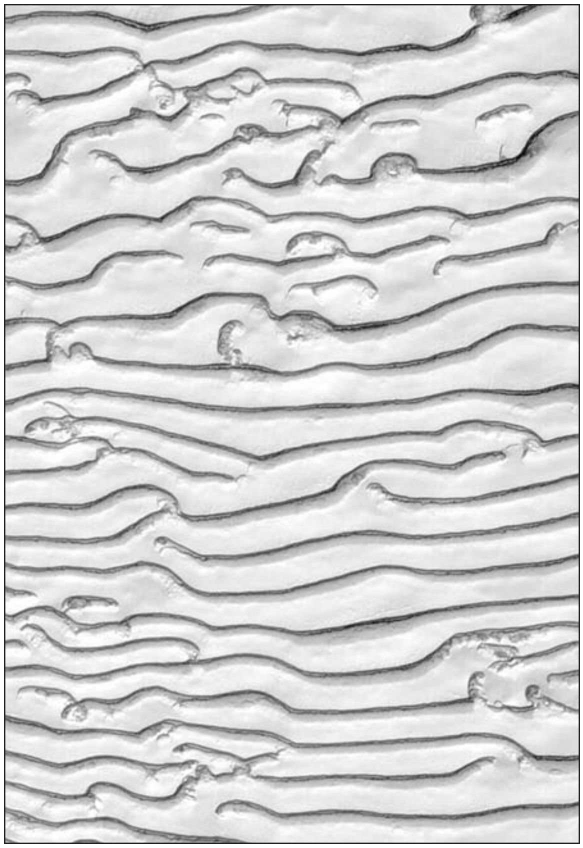
Mars Express was launched by the European Space Agency (ESA) on June 2, 2003. It represented the first planetary mission attempted by ESA and was followed a few years later by a similarly designed probe sent into Venus orbit, the Venus Express. Mars Express entered Mars orbit on December 25, 2003. It carried a radar system capable of detecting subsurface water ice in permafrost regions. In a near-polar orbit, it was able to study surface changes and make mineralogical maps in addition to monitoring sea-



Spiderweb patterns emerge after the seasonal disappearance of the Martian north pole's carbon-dioxide ice cap. (NASA/JPL/Malin Space Science Systems) sonal changes in the polar caps. Mars Express's Fourier spectrometer discovered that methane was entering the atmosphere in near-equatorial regions having subsurface ice. Its Visible and Infrared Mineralogical Mapping Spectrometer detailed hydrated sulfates. Radar data provided evidence of underground water ice. Mars Express found the signature of water ice in the planet's south pole. What some interpreted as a deposit of water ice was discovered on the floor of an otherwise nondescript crater in the vicinity of Mars's north pole.

Mars Reconnaissance Orbiter (MRO) was launched on August 12, 2005, and injected into orbit about Mars on March 10, 2006, joining five other operational spacecraft already performing scientific investigations of the Red Planet. MRO was outfitted with the largest imaging system (HiRISE) ever sent to another planet, as well as spectrometers and a penetrating radar. Among its primary science objectives were searches for water in the polar regions and a daily determination of meteorological conditions that might lead to changes in the state of Mars's water deposits. HiRISE images provided evidence of banded terrain, suggesting the action of water on the surface in relatively recent geological times. The spacecraft's radar was able to detect underground deposits of water ice.

Mars Phoenix successfully touched down on Mars's northern arctic region on May 25, 2008, in an area largely devoid of large and medium-sized rocks. Its landing marked the first time since the Viking spacecraft that a powered lander had successfully reached the Martian surface. The landing site was chosen to have a high possibility of a layer of ice either on the surface or not very deep beneath. Mars Phoenix was outfitted with a robotic arm, at the end of which a scoop was attached. The scoop not only could pick up loose surface soils but also could scrape the ice. The idea was to lift soil and potential ice into a special gas analyzer that would process the soil and chemically analyze it for the presence of water and organic molecules. Very soon after Mars Phoenix landed, initial photographs strongly hinted that the lander had touched down near ice. Under the lander it appeared that exhaust from Mars Phoenix's rockets had kicked up the soil, exposing subsurface ice. Early attempts to analyze the soil were thwarted by communications problems, and the entryway into the gas analyzer was clogged by soil that clumped. The very first run of the lander's Thermal Evolution and Gas Analyzer produced a disappointing lack of any water signature. However, subsequent data did provide strong evidence of water at the lander's site.



The Martian south pole's characteristic "swiss cheese" terrain, carved into the carbon-dioxide ice cap. (NASA/JPL/Malin Space Science Systems)

Also in May, 2008, scientists studying MRO data revealed that it appeared that Mars's crust and upper mantle could be even colder than originally believed. If true, that would mean that liquid water would have to be found even deeper, where internal heat would permit water to exist in that phase. MRO's Shallow Radar experiment probed the internal structure of ice, sand, and dust layers on the north polar cap and revealed the planet's lithosphere (combination of crust and mantle) to be thicker and colder than previous models indicated, and as such provided greater support against the stress of the weight of the polar cap atop it. Radar imagery indicated four zones of thin ice and dust layers interspersed with rather thick layers of water ice going down into the lithosphere to a geological depth such that it recorded changes in climate on Mars over many millions of years. The thin layers represented climate changes perhaps lasting only one million years. These data did not shed light on the cause of the climate change, but two strong possibilities were alterations of Mars's rotational axis and/or orbital eccentricity.

CONTEXT

The question whether there is now or could ever be life on Mars has long been a matter of conjecture. If there is, was, or will be life on Mars, that life would likely be confined to organisms much smaller and less complex than anything resembling human beings. To sustain animal or vegetable life, it is presumed, some minimal atmosphere must exist and water must be present. Research into the polar caps of Mars provides evidence that water is plentiful in the planet's frozen polar caps. The same research suggests that the planet is much drier than Earth, that its atmosphere is so rarefied that it cannot support complex organisms, and that its temperatures (although much less forbidding than those on Jupiter or Venus) are not conducive to life.

In the second quarter of the twentieth century, Eugene M. Antoniadi, a Greek-born astronomer who spent his professional life in France, explained the dark areas on the periphery of the polar caps in summer, designated the "Lowell bands" for astronomer Percival Lowell. Antoniadi thought that the fringes of the ice fields were reduced in brightness by grasses and bushes that grew in the periphery, presupposing by such a contention the existence of life on Mars and of an atmosphere that would support life. This theory has been disproved, but it reflected the widespread notion that life exists on Mars—although not in the form of the "little green men" that some works of science fiction describe.

Spacecraft exploration of the polar caps of Mars has revealed that water in some form exists on Mars and has suggested that water in liquid form was once more plentiful there than it is currently. The channels of Mars are now generally thought to have been forged through the eons by flowing water, and the presence of the polar ice caps supports this theory. These explorations have also presented evidence that the atmosphere of Mars has changed drastically from what it once was. Climatic change may one day produce a Mars quite different from the present planet. Astronomers who make such projections, however, caution that they are talking about millions, perhaps hundreds of millions, of years. Mars is unlikely to change in easily perceptible ways within the life span of a single human now living.

In the aftermath of the 2003 space shuttle *Columbia* accident, President George W. Bush redirected NASA to complete the International Space Station and retire the shuttle fleet by 2010 in order to move on to human exploration beyond low-Earth orbit. Project Constellation was intended to take astronauts back to the Moon by 2020 to begin a permanent research outpost there and in time to send human explorers to Mars. In order to make both of those goals happen, the use of local water would be required. Sending humans to Mars would be possible only if the robotic search for water on Mars bore fruit; the polar ice caps would play prominent roles in liberating water from the surface of Mars for use by future astronauts.

R. Baird Shuman

FURTHER READING

Barlow, Nadine. *Mars: An Introduction to Its Interior, Surface, and Atmosphere*. Cambridge, England: Cambridge University Press, 2008. An interdisciplinary text including contemporary data from the Mars Exploration Rovers and Mars Express. A good reference for planetary science students and nonspecialists alike. Each chapter contains necessary background information.

Carr, Michael H. *The Surface of Mars*. Cambridge, England: Cambridge University Press, 2007. The author provides a complete description of the geological heating of Mars as understood based on results from Mars Global Surveyor, Mars Odyssey, Mars Reconnaissance Orbiter, Mars Express, and the Mars Pathfinder and Mars Exploration Rovers. Heavily illustrated; includes a comprehensive reference list.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that have investigated all planets with the exception of Pluto by the publication date. Takes an astrophysical approach to give our solar system a wider context as just one member of similar systems throughout the universe.

Glasstone, Samuel. *The Book of Mars*. Washington, D.C.: National Aeronautics and Space Administration, 1968. This classic work contains a useful history of Mars up through the early Mariner flyby missions. Provides excellent illustrations. Accessible to laypersons.

Harland, David M. *Water and the Search for Life on Mars*. New York: Springer Praxis, 2005. A historical review of telescope and spacecraft observations of the Red Planet up through the Spirit and Opportunity rovers. Covers all aspects of Mars exploration, but focuses on the search for water, believed to be the most necessary ingredient for life.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science. The material on Mars covers all aspects of ground-based and spacecraft observations of Mars. Takes a comparative planetology approach rather than presenting individual chapters on each major object in the solar system.

Hartmann, William K., and Odell Raper. *The New Mars: The Discoveries of Mariner 9*. Washington, D.C.: National Aeronautics and Space Administration, 1974. Although its information was rendered somewhat outdated by the Viking missions, this book is valuable for its detailed information about the material uncovered by Mariner 9. Includes numerous photographs; they will be useful for students of the polar caps, because Mariner 9 photographed the south polar cap steadily for four months. Some background in astronomy would be helpful.

Kargel, Jeffrey S. *Mars: A Warmer, Wetter Planet*. New York: Springer Praxis, 2004. A member of Springer Praxis's excellent Space Exploration series. The author provides a convincing case that the picture of a dry, waterless world portrayed initially by the early Mariner probes is not the Mars of today's understanding. The book takes the reader on a search for Mars's water.

Lowell, Percival H. *Mars and Its Canals*. New York: Macmillan, 1906. Lowell speculates that the polar caps are composed of carbon dioxide, a theory that regained its vogue after Mariner 4's pictures suggested an absence of water in substantial quantities in the polar caps. The book has historical interest for sophisticated readers.

Moore, Patrick. *Guide to Mars*. New York: W. W. Norton, 1977. Dated, but a reliable book on Mars and Mars exploration, this compact volume provides an excellent chapter that focuses on the Martian ice caps, plains, and deserts. The material on the Mariner and Viking missions and their results is indispensable to any serious student of the polar caps or of Mars in general. Necessary historical information is skillfully woven into the text.

Richardson, Robert S. *Exploring Mars*. New York: McGraw-Hill, 1954. It is interesting to compare this dated book with later ones that clearly indicate the enormous progress that has been made in understanding Mars. The material on the polar caps is limited but valuable. The index is extensive and accurate.

Squyres, Steve. *Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet*. New York: Hyperion, 2006. Written by the principal investigator for the Mars Exploration Rovers Spirit and Opportunity, this fascinating book provides a general audience with a behind-the-scenes look at how robotic missions to the planets are planned, funded, developed, and flown. A personal story of excitement, frustrations, a scientist's life during a mission, the satisfaction of overcoming difficulties, and the ongoing thrills of discovery.

See also: Earth-Sun Relations; Ganymede; Mars: Illusions and Conspiracies; Mars: Possible Life; Mars: Surface Experiments; Mars's Atmosphere; Mars's Craters; Mars's Satellites; Mars's Valleys; Mars's Volcanoes; Mars's Water; Meteoroids from the Moon and Mars; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Pluto and Charon; Solar Variability; Triton.

Mars's Satellites

Categories: Mars; Natural Planetary Satellites; Planets and Planetology

The two satellites of Mars, Phobos and Deimos, almost certainly originated as captured asteroids. These two small satellites could serve as future way stations for human exploration of Mars. They may provide future astronauts visiting Mars with an orbital base and even essential resources.

OVERVIEW

In early August, 1877, astronomer Asaph Hall began his search for Martian satellites at the U.S. Naval Observatory in Washington, D.C. His search was initiated for two primary reasons. He found that many astronomy texts and ephemerides of the day contained serious errors and incorrect statements. One contention was that Mars had no satellites; because none had yet been identified, that statement was correct for its time. However, Hall knew from consulting Frederick Kaiser's summary of Martian observations in the *Annals of the Leiden Observatory* (1872) that few astronomers had even looked for such bodies. While Mars made its close approach to the Earth (called "opposition") in 1877, Hall used the Naval Observatory's 26-inch Clark refractor telescope to search for potential Martian satellites.

Even as he began his search, Hall knew that the probability of finding a Martian satellite was slim. Any object even a fraction of the size of Earth's moon would have been discovered long before. Any smaller object could not even exist at any great distance from Mars, as the Sun's gravitational influence would snatch it away. Hall therefore began his search looking for a very small satellite orbiting very close to the planet, one that therefore might be very close to the visible disk of Mars as seen through the telescopic lens and thus obscured in the planet's glare. In view of these discouraging considerations, Hall said, "I might have abandoned the search had it not been for the encouragement of my wife." This statement would become indelibly etched on the discovery. On August 12, 1877, Hall first glimpsed one of Mars's two satellites, which he confirmed on August 16. The next evening, he discovered a second satellite. The announcement was made several days later by Navy admiral John Rodgers, the observatory's superintendent. These bodies were named Deimos (meaning "flight" or "panic") and Phobos ("fear") from Homer's *Iliad*: "He [Mars] spake, and summoned Fear and Flight to yoke his steed, and put his glorious armor on."

The first clear images of the satellites were made by the National Aeronautics and Space Administration's (NASA's) Mariner 9 orbiter in 1972. Five years later, even more dramatic and detailed photographs were obtained by the Viking orbiters. Those photographs revealed that the two satellites are among the darkest bodies in the solar system. Because of their density, size, and curious orbital characteristics, they are widely thought to be asteroids captured by Mars's gravitational field. They are not circular in shape. Some have even described their appearance as potato-shaped; their modest size allows for the weakest of gravitational fields, which will not permit the body to collapse into a spherical shape like that of planets and much larger satellites.

Phobos, the larger of the two Martian satellites, has a size of 28 by 23 by 20 kilometers. Phobos's orbit is exceptionally low—directly above the equator only 9,378 kilometers over the planet's surface—so low that it cannot be observed from the surface at latitudes greater than 70° north or south. The orbit of Phobos is just barely outside the Roche limit, where tidal forces would tear it apart. It probably will be torn apart and crash into the surface of Mars in the next 38 million years. Phobos's orbital period is very fast. It circles the planet in only 7 hours and 39 minutes, making the satellite appear to rise in the west and set in the east.

Deimos measures 16 by 12 by 10 kilometers. Its orbit is considerably higher than that of Phobos: 23,459 kilometers above the equator, with a period of 30 hours and 18 minutes. Deimos's orbit is high enough that it should eventually escape Mars's gravitational field and fly off into an independent orbit around the Sun.

Initial observations of the two satellites showed them to be heavily cratered, suggesting that their surfaces are very old—nearly as old as the solar system itself, some 4.6 billion years. There are two very large craters (relative to the body's size) on Phobos. The largest crater, dominating Phobos's northern hemisphere, is named Stickney, for discoverer Hall's wife. The astronomer himself has been remembered by a lesser surface marking: Crater Hall is a 6-kilometer depression on the satellite's southern pole. Crater Roche lies near Phobos's north pole, a reminder of the satellite's eventual fate as a result of Mars's dominant gravitational influence. As a Viking orbiter flew to within 88 kilometers of Phobos's surface, it photographed what appeared to be cracks, as wide as several hundred meters and up to 10 kilometers in length, emanating from Crater Stickney. These cracks in the satellite's surface were almost certainly caused by the impacting body that formed the crater itself.

The surface of Deimos is not as spectacular as that of the larger satellite. A Viking orbiter flew to within 23 kilometers of Deimos and resolved a relatively quiescent surface, with smaller craters (the largest discovered is only 2.3 kilometers across), no visible cracks, and a lack of a single spectacular formation.

This image of Phobos, the larger of Mars's two moons, was captured by the Mars Reconnaissance Orbiter's High Resolution Imaging Science Experiment (HiRISE) on March 23, 2008. (NASA/JPL-Caltech/University of Arizona)

The theory of asteroid origins was bolstered by these visual images. An extraordinarily dark surface and a density only twice that of water (half that of Mars) lent credibility to the theory that the satellites were composed of carbon and carbon compounds such as are conjectured for type C asteroids, which populate the outermost regions of the asteroid belt. These bodies may contain up to 20 percent water by weight. Why their orbits are so nearly perfectly circular and equatorial has no easy explanation. In terms of capture probability, wayward asteroids would not necessarily slip into such neat orbits; the resulting orbit would most likely be inclined and considerably eccentric. Hence, for confirmation, such theories will have to await both the physical exploration of their surfaces and a firsthand look at carbonaceous asteroids themselves.

The Soviet Union's probes to the Martian satellite's surface were collectively called the Phobos mission. Two craft, Phobos 1 and Phobos 2, were launched in the summer of 1988; their mission was to send a lander to Phobos so that it could analyze the satellite's surface. Communication with Phobos 1 was lost before it reached Mars's orbit. Phobos 2 successfully attained orbit around Mars in 1989 and began returning photographs of Phobos and Mars. On March 27, 1989, merely days before the probe's planned close encounter with Phobos and release of its landers, the spacecraft spun irretrievably out of control, and the mission was lost.

Some Facts About Mars's Satellites

	Phobos	*Deimos*
Semimajor axis* (km)	9,378	23,459
Sidereal orbit period (days)	0.31891	1.26244
Sidereal rotation period (days)	0.31891	1.26244
Orbital inclination (deg)	1.08	1.79
Orbital eccentricity	0.0151	0.0005
Major axis radius (km)	13.4	7.5
Median axis radius (km)	11.2	6.1
Minor axis radius (km)	9.2	5.2
Mass (10^{15} kg)	10.6	2.4
Mean density (kg/m^3)	1,900	1,750
Geometric albedo	0.07	0.08
Visual magnitude V(1,0)	+11.8	+12.89
Apparent visual magnitude (V0)	11.3	12.40

*Mean orbital distance from the center of Mars.
Source: National Space Science Data Center, NASA Goddard Space Flight Center.

Intense interest in the tiny satellites of Mars has been generated for several reasons. They provide a natural "space station" for future Mars explorers. Their tiny gravitational fields require very little energy to overcome, but they offer a stable platform for the staging of expeditionary landing and observation teams. They may also contain substantial quantities of water that may one day be mined to provide hydrogen and oxygen for space travelers, thus reducing the necessary burden of transporting it from Earth to the surface of Mars. Finally, less energy is required for a mission to the satellites of Mars than to and from the surface of the Earth's moon because of their weak gravitational fields and hence their low escape velocities. Such a mission, which would require a minimal round-trip travel time of two to three years, was seriously discussed by both the Soviet government and certain interests in the United States near the time when the Soviet Union collapsed in 1991. Two decades then passed with little interest in Phobos or Deimos on the part of the two major spacefaring nations.

METHODS OF STUDY

NASA spacecraft that encountered Phobos and Deimos produced photographic studies from onboard cameras and mass studies based on flyby navigational data. The most detailed photographic studies were conducted with Viking orbiter cameras. Viking 1 flew by Phobos on February 12, 1977. Viking 2 flew by Deimos on September 25, 1977. The photographic system on the orbiters was called the Viking Imaging System. It returned a total of 51,539 images of Mars and its satellites to Earth. The masses of Phobos and Deimos were estimated by determining how much the Viking spacecraft were deflected in their orbits around Mars by the gravitational fields produced by the satellites. Until the Viking encounters, the masses of Phobos and Deimos were not precisely known. Through observation of their orbital motion, coupled with a Martian mass determination, these encounters provided good estimates of the masses of Phobos and Deimos.

Photographic studies alone produced a wealth of information. A technique known as reflectivity—measuring the amount of light reflected from the surface of the satellites—enabled planetary scientists to speculate that the satellites may have been captured asteroids. It has long been speculated that a class of asteroids made of carbon and carbon compounds would be exceedingly dark, as Phobos and Deimos proved to be.

The surfaces of both satellites are saturated with craters, which indicates that their surfaces are very old. This finding enabled dating of the asteroids. In addition, the peculiar 10-kilometer crater on Phobos named Stickney displayed very large cracks down the surface of the satellite, which hinted to some planetary geologists that the composition of the body may contain substantial amounts of water ice. The cracks also indicated other subsurface structural features as well as the depth of the regolith, or top layer of soil overlying the bedrock.

CONTEXT

Phobos and Deimos may someday become two of the most important way stations in the solar system. As the Earth's focus of exploration turns to Mars as the next most logical frontier of exploration and colonization beyond the Moon, Phobos and Deimos will serve as stepping-off points to the surface of Mars. They could provide a base for operations to and from the planet. They could also provide a communications base for ground-to-space and -Earth information exchanges. Finally, these two tiny satellites might be able to supply water, fuel, and oxygen to future space explorers.

Plans have been considered for sending a crewed mission to Phobos as a dress rehearsal for a later mission to the surface of Mars. This mission would test the critical life-support and medical issues that currently limit a Mars mission, at a fraction of the energy that would be required to land on the surface of the planet itself.

In the aftermath of the *Columbia* accident (February 1, 2003), the Bush administration reconsidered NASA's mission. In early 2004 the White House directed NASA to complete the International Space Station and retire the space shuttle fleet by 2010, and then to proceed with human exploration beyond low-Earth orbit. This plan called for a return to the Moon to stay and then an evolutionary approach leading to a crewed expedition to Mars. Phobos and Deimos will undoubtedly play a prominent role in eventual plans for the human exploration of Mars.

Dennis Chamberland

FURTHER READING

Collins, Michael. *Mission to Mars*. New York: Grove Weidenfeld, 1990. Apollo 11 astronaut Collins provides an astronaut's vision of a trip to Mars. Examines the problems to be overcome to make such a journey possible using space-shuttle-era technology.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that had investigated all planets with the exception of Pluto by the publication date. Takes an astrophysical approach, considering the solar system as just one member of similar systems throughout the universe.

Ezell, Edward, and Linda Ezell. *On Mars: Exploration of the Red Planet, 1958-1978*. NASA SP-4212. Washington, D.C.: Government Printing Office, 1984. This classic book is the National Aeronautics and Space Administration's official history of the Viking program, from the original ideas in 1958 to the culmination of the project some twenty years later. Includes discussion of the Martian moons. Generally nontechnical and accessible to all readers.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science, covering all aspects of ground-based and spacecraft observations of Mars. Takes a comparative planetology approach rather than including separate chapters on each planet in the solar system.

Hartmann, William K., et al. *Out of the Cradle: Exploring the Frontiers Beyond Earth*. New York: Workman, 1984. This combination picture book and narrative of future human exploration discusses Mars as a logical next step for human exploration and settlement after the Moon. Depicts the potential struggles of future colonists on Mars and includes a discussion of the Martian satellites. Illustrated with an artist's conceptions of future planetary bases and explorers.

Jöels, Kerry Mark. *The Mars One Crew Manual*. New York: Ballantine Books, 1985. This speculative facsimile of a future explorer's crew manual offers an excellent reference for what the future colonist may find on arrival at Mars. Dated, but written for all backgrounds; heavily illustrated.

Kargel, Jeffrey S. *Mars: A Warmer, Wetter Planet*. New York: Springer Praxis, 2004. A member of Springer Praxis's excellent Space Exploration series. The author provides a convincing case that the picture of a dry, waterless world portrayed initially by the early Mariner probes is not the Mars of today's understanding. The book takes the reader on a search for Mars's water.

Mutch, Thomas A., et al. *The Geology of Mars*. Princeton, N.J.: Princeton University Press, 1976. A complete geological analysis of the recovered Viking data, accessible to those with a college-level science background. Illustrated with photographs and tables.

National Aeronautics and Space Administration. *The Case for Mars: Concept Development for a Mars Research Station*. San Francisco: University Press of the Pacific, 2002. This is a speculative study of how a Mars base could be established. Ironically, it came out two years before NASA was given the task of developing an evolutionary human exploration program to leave low-Earth orbit, return to the Moon to stay, and eventually travel to Mars.

Sheehan, William, and Stephen James O'Meara. *Mars: The Lure of the Red Planet*. Amherst, N.Y.: Prometheus Books, 2001. Takes a different approach to the investigation of Mars, examining what makes the Red Planet so alluring. Also describes the great astronomers who advanced humanity's understanding of Mars, from ancient times to the space age.

Zubrin, Robert. *Entering Space: Creating a Spacefaring Civilization*. New York: Tarcher, 2000. The author displays a gung-ho attitude toward making humanity a truly spacefaring species by accepting the challenge of journeying to Mars sooner rather than later, using contemporary technology and daring innovation. Speculates beyond travel to Mars.

See also: Mars: Illusions and Conspiracies; Mars: Possible Life; Mars: Surface Experiments; Mars's Atmosphere; Mars's Craters; Mars's Polar Caps; Mars's Valleys; Mars's Volcanoes; Mars's Water; Meteoroids from the Moon and Mars; Planetary Satellites; Saturn's Satellites.

Mars's Valleys

Categories: Mars; Planets and Planetology

Like Earth, Mars has valleys exhibiting complex geological histories, including flowing water, hill-slope processes, and structural control. Unlike Earth, yet similar to the Moon, Martian valleys may be as old as 4 billion years.

OVERVIEW

Valleys are low-lying, elongate troughs on planetary surfaces that are surrounded by elevated ground. On Earth, valleys often contain a stream with an outlet. About two hundred years ago, the origin of valleys on Earth was very controversial. In 1788, the Scottish naturalist James Hutton disputed the prevailing opinion that valleys formed by cataclysmic flooding, specifically the Noachian flood of biblical accounts. Hutton hypothesized that valleys formed gradually through the erosive action of the rivers and streams that lay on their floors. The fluvial origin of most valleys was subsequently demonstrated by detailed geomorphological work over the next century.

On the Moon, which lacks significant water and other volatile chemical components, valleys are very different. Valleys on the Moon are completely dry and are thought to have been created mostly by subsurface forces. Some lunar valleys may also form by chains of craters left by impacting meteors. Some, called sinuous rilles, have formed by the erosive action of lava. Others are structural depressions formed as surface blocks dropped between fractures. Such valleys also occur on Mars and Earth, but they are much less common than the fluvial forms.

In 1972, spacecraft images revealed the presence of channels and valleys on Mars that appeared very similar to those on Earth. On Mars, a very interesting inversion of scale occurs for channels and valleys. Channels are those troughs in which fluid flow once completely surrounded the depression that confined it. Martian channels are up to 200 kilometers wide and 2,000 kilometers long. The channels are much larger than networks of small valleys in the ancient, heavily cratered uplands of the planet.

Ius Chasma's southern trench, near the Valles Marineris—the "Grand Canyon" of Mars. Layers of lava flows alternate with bright layers of sedimentary rock, which may be the alluvia carried by ancient Martian water flowing on the planet's surface. (NASA/JPL-Caltech/University of Arizona)

The valleys are typically a few kilometers wide and several hundred kilometers long. Both Martian channels and Martian valleys are extremely ancient by terrestrial standards but are comparable in age to many features on the Moon. The valleys formed early in the planet's history—by analogy to the Moon, more than 3.5 billion years ago—when rates of impact cratering were much higher than they are at present. The channels are somewhat younger, extending in age from about 3 to 0.5 billion years ago.

Martian channels were formed by immense flows of fluid that emanated from zones of collapsed topography known as chaotic terrain. This fluid seems to have burst onto the surface as immense floods of water plus considerable sediment. In the extreme cold of the Martian environment, the water would have partly frozen to form ice jams driven by the turbulent water. Local blockades of ice may have induced secondary floods, and the ice itself could have flowed in a manner somewhat similar to terrestrial glaciers. Processes of cataclysmic water outburst were probably repeated over long periods of time. These floods were probably generated from ground ice in the subsurface, perhaps heated by volcanic activity.

One of the enigmas about flowing water on Mars is the present surface environment of the planet. Surface atmospheric pressure is only about 0.7 percent of terrestrial atmospheric pressure at sea level. The temperatures measured at the Viking landing sites on the planet ranged from 243 kelvins by day to 193 kelvins at night. Subsequent landers verified the surface conditions first reported by the Vikings. Under these conditions, any standing body of water would rapidly vaporize and freeze. Rapid outbursts of water that formed the large channels, however, could have been maintained because of the relatively short duration of the flow events.

The existence of small valley networks of the heavily cratered terrains of Mars poses a problem for scientists. The valleys have short tributaries that end in abrupt valley heads, similar to the box canyons of the western United States. This valley form is believed to be caused by spring sapping. Sapping is the process whereby groundwater undercuts hill slopes; the groundwater apparently emerged as springs at the heads of the valleys, providing a subsurface source for flow.

One way to have maintained the groundwater flow that sustained Martian valley growth would have been for precipitation (rain or snow) to have fallen in the headwater areas. Water infiltrating the ground would then have recharged the groundwater flow system. This mechanism, however, requires Mars to have once had a very different climate from the one observed today. The atmosphere would have had to be warm and dense enough to hold considerable water. Atmospheric scientists have constructed theoretical models of such an ancient, hypothetical atmosphere for Mars. They conclude that increased amounts of carbon dioxide may have been present early in the planet's history. The carbon dioxide could have contributed to a greenhouse effect, whereby the planetary surface is warmed by trapping incoming solar radiation as heat. An alternative mechanism for maintaining groundwater flow to the valley networks would have been for geothermal systems to drive the flow by convective, or circulating, heat. Subsurface volcanic rocks would supply the heat, circulating the groundwater in a manner similar to that in areas of hot springs, such as Yellowstone National Park. Water flowing in the valleys would cool, seep into the ground, and recharge a recirculating system driven by volcanic heat. This mechanism would not require a dense, warm atmosphere early in the planet's history.

Both channels and valleys on Mars are modified by many other processes besides water flow. Because these features have hill slopes, a variety of gravity-induced slope adjustment forms are present; called mass movements, these slope adjustments include landslides, flows of debris, and slow creep of slope materials. All these processes may have been facilitated by water and ice mixed with the rock materials. Movement of debris onto the channel and valley floors, in some cases, completely conceals evidence of fluvial action that originally cut into the landscape. The Mars Global Surveyor spacecraft provided images, taken years apart, that demonstrated slumping of walls, albeit in craters, that likely was the result of water under the surface diminishing the load-bearing capability of those walls.

Wind action is also facilitated by the confinement of valleys. Wind erodes fine sediment, producing a lineated topography that parallels prevailing directions. Eroded materials may locally accumulate as sets of sand dunes, or they may be more broadly distributed as sheets of deposited sand or dust.

Valleys also served as troughs along which erupted lavas descended from volcanic source areas. Indeed, Martian volcanoes serve as excellent sites in which to observe the evolutionary sequence of valley development on Mars. Volcanoes vary in age and in the character of their surfaces, thereby providing a kind of natural experiment on the formation of valleys. Studies of fluvial valley development on volcanoes indicate that incision by flowing water occurs only when the very permeable lava flows of the volcanoes are altered to have less permeable surfaces. Lowered surface permeability arises from volcanic ash that mantles local areas. Channels forming on this ash incise into the volcano. As valleys form by enlargement of these zones, the incision is able to tap groundwater in the permeable lava flows. This groundwater further sustains valley growth in a headward direction by sapping. Eventually, the volcano is dissected by a mature network of valleys adjusted to the water flow system that is sustaining their growth.

Many mysteries still surround the valleys on Mars despite a rich history of spacecraft imaging over four decades from Mariner 4 to the Mars Reconnaissance Orbiter and Mars Express. While most valley networks are very old—older than most rocks on Earth—some are quite young. Very well developed valleys occur on the relatively young Martian volcano Alba Patera. Valleys there are restricted to a local area of volcanic ash. It may be that water was introduced by local precipitation, perhaps related to outburst flooding in the large channels.

Mars has additional surprises, including an abundance of linear grooves and a lack of depositional forms. Also puzzling is the fact that the large channels on Mars contain landforms that are very different from landforms generally seen on Earth. The best Earth analogy to the outflow channels is a region called the Channeled Scabland in Washington State. This area of flood-eroded basalt was generated by immense glacial-lake outbursts during the ice ages. The Channeled Scabland is more similar to the Martian channels than any other region on Earth.

METHODS OF STUDY

Channels and valleys of Mars were discovered by remote-sensing observations generated from spacecraft. Despite Percival Lowell's accounts of Martian "canals" nearly one hundred years ago, telescopic views of the planet were inadequate to interpret the presence of channels and valleys. It was not until 1972, when the Mariner 9 spacecraft returned the first high-resolution pictures of them from orbit about

Mars, that the importance of the valleys was realized.

Ages of the valleys are interpreted by the numbers of superimposed impact craters. Then, by analogy to known cratering rates and histories on the Moon, ages can be assigned to the Martian landforms. The genesis of the valleys must also be determined by analogy. An interpreter of the pictures of the valleys must have a broad familiarity with natural landscapes, which is used to infer causes for the combinations of features seen in the channels and valleys. Often details of planetary landforms are somewhat different from what is generally known from terrestrial experience. On Mars, such lack of correspondence arises from lower gravitational acceleration, the low surface pressure, and low temperature.

Details of Martian landforms are analyzed on special maps that show relationships and patterns. Geological maps elucidate the time sequence of development, and geomorphological maps show relationships among the planetary features. Quantitative measurements can be made of the landform shapes, which can be compared to measurements on similar-appearing terrestrial features.

Model building is the activity whereby an explanation of the valley is provided in a form that extracts significant elements from the natural complexity of the phenomenon. The model may be expressed in abstract mathematical terms; it can involve laboratory hardware; or it can simulate the sequence of landform development through various kinds of analogy. In all cases, however, the model is used to express in simple, predictable terms the complexity of the real-world system under investigation. Models are only as good as their correspondence to the natural system, however, so successful model building requires an intimate knowledge of the system under investigation. This knowledge must be continually checked against new data about that system gained through ongoing investigation.

Mars studies can be separated into two distinct categories. During the initial discovery phase, Mars was visited by Mariner spacecraft. Mariner 4 flew by, snapping a small number of pictures in 1965 that revealed a Moon-like character to the surface of the Red Planet. It was something of a disappointment that the more sophisticated Mariner 6 and 7 flyby missions in 1969 verified the arid, cratered character indicated by the previous Mariner spacecraft. Then Mariner 9 orbited Mars and revealed valleys and river features as well as giant volcanoes, indicating that Mars had a dynamic past. The Viking landers examined the plains and found ambiguous soil analysis results that could not definitely answer the question about life on Mars. These missions completed the discovery phase.

Beginning with the failed Russian Phobos missions and followed by the National Aeronautics and Space Administration's (NASA's) robust sequence of spacecraft—Mars Observer (unsuccessful), Mars Pathfinder (a rover), Mars Global Surveyor, Mars Climate Orbiter (unsuccessful), Mars Polar Lander (unsuccessful), Mars Odyssey, the Mars Reconnaissance Rovers, and Mars Reconnaissance Orbiter—along with the European Space Agency's Mars Express, a second phase of focused Mars studies examined planetary features in large part seeking an answer to one fundamental question in particular. That question was essentially where Mars's water had gone.

A third phase of Mars exploration will begin once humans develop a capability to journey to Mars and complete research in situ and travel across the Martian surface with sophisticated instruments.

CONTEXT

The valleys of Mars reflect immense environmental changes that occurred on the planet. Surface conditions presently are too cold and the atmosphere too thin for water to exist in its liquid state. However, in the past, during selected epochs of its planetary history, Mars seems to have been able to sustain an active hydrological cycle that produced river valleys similar to those on Earth. Because the small valleys occur throughout the heavily cratered terrains of the planet, the scale of environmental change must have been global. In July, 2008, the Mars Phoenix lander sampled subsurface water ice near the northern polar region, providing the first direct evidence of water still on Mars.

In 2008 data from the Mars Reconnaissance Orbiter led to the conclusion that Mars of the past possessed a wet environment, something that had been suspected after the tantalizing images from Mariner 9 led scientists to alter their assessment of the planet Mars after the disappointing dry, cratered surface revealed by the Mariner 4, 6, and 7 flyby missions. Current studies strongly indicated that ancient highlands on Mars, which constitute nearly half the planet's surface, contain certain types of clay minerals that need water in order to form. This suggests that Mars once had large lakes, flowing rivers, and smaller wet environments for prolonged periods of time in the distant past. Water erosion moved the clay minerals into delta formations. The crater Jezero appears to have once confined a lake that was later breached; its water carried the clay minerals out into a fan-shaped structure.

Earth has also been affected by global environmental change. Numerous times over the past several million years, the planet has experienced decreased global temperatures with associated glaciation. During glacial advances, hydrological conditions were profoundly changed. Most recently, the global change of interest on Earth has become the warming associated with artificially increased levels of carbon dioxide and other trace gases in the atmosphere. Thus, like Mars, Earth oscillates between periods of increased warmth and cold. By comparing theories that explain such cycles, scientists hope to understand exactly how the environmental change occurs and thus to develop a means of predicting future change that will affect humanity on its home planet.

In a broad sense, the development of humankind has been linked to discoveries associated with exploration and with the migration of peoples to new lands; the space program is the most modern manifestation of such trends. When the channels and valleys of Mars were discovered, they stimulated an immense scientific effort to explain the conditions on Mars that made flowing water possible in the past. Scientists generally agree that most of the water on Mars is now locked up in its subsurface, frozen as ground ice in layers of thick permafrost. There is speculation that, if Mars could be made warm, perhaps by inducing plant growth or by releasing trapped carbon dioxide, it is not inconceivable that Martian rivers could be made to flow again—and might even sustain a population of emigrants from Earth.

Victor R. Baker

FURTHER READING

Atreya, S. K., J. B. Pollack, and M. S. Matthus, eds. *Origins and Evolution of Planetary and Satellite Atmospheres*. Phoenix: University of Arizona Press, 1989. A collection of research articles by planetary scientists. Provides a comparative look at Venus, Earth, and Mars.

Baker, Victor R. *The Channels of Mars*. Austin: University of Texas Press, 1982. This 198-page book provides a complete review of scientific ideas about channels and valleys on Mars. Abundantly illustrated with pictures of Martian landforms, diagrams illustrating processes, and interpretive maps. Dated, yet provides considerable discussion of possible terrestrial analogues to features on Mars. Also provides general review material on the history of Mars studies, the general geology of Mars, and implications for global environmental change on the planet.

Barlow, Nadine. *Mars: An Introduction to Its Interior, Surface, and Atmosphere*. Cambridge, England: Cambridge University Press, 2008. An interdisciplinary text including contemporary data from the Mars Exploration Rovers and Mars Express. Each chapter contains necessary background information. Excellent for both students and nonspecialists.

Carr, Michael H. *The Surface of Mars*. New Haven, Conn.: Yale University Press, 1981. A well-illustrated summary of Mars geology, this book extensively features pictures from the Viking orbiter spacecraft. One chapter covers channels and valleys, and all the important landforms of Mars are discussed in some detail. This classic 232-page book is the best starting point for an in-depth review of modern ideas about the geology of Mars.

Greeley, Ronald. *Planetary Landscapes*. Winchester, Mass.: Allen & Unwin, 1994. Provides a review of the geomorphology of planetary surfaces throughout the solar system.

Compares the processes among multiple planetary bodies with an emphasis on properties of individual planets and moons. Channels and valleys on Mars are seen in the broader context of surfaces on the various planets.

Harland, David M. *Water and the Search for Life on Mars*. New York: Springer Praxis, 2005. A historical review of telescopic and spacecraft observations of the Red Planet up through the Spirit and Opportunity rovers. Covers all aspects of Mars exploration, but focuses on the search for water, believed to be the most necessary ingredient for life.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science. The material on Mars covers all aspects of ground-based and spacecraft observations of Mars. That information appears in several chapters, as the text takes a comparative planetology approach.

Kargel, Jeffrey S. *Mars: A Warmer, Wetter Planet*. New York: Springer Praxis, 2004. A member of Springer Praxis's excellent Space Exploration series, arguing that the dry, waterless world portrayed initially by the early Mariner probes is not the Mars of today's understanding. The book takes the reader on a search for Mars's water.

Mutch, Thomas A., et al. *The Geology of Mars*. Princeton, N.J.: Princeton University Press, 1976. This classic 400-page book was produced at the beginning of the Viking mission to Mars as a summary of knowledge up to that time. Reflects the discovery of Martian channels by the Mariner 9 mission in 1972. Although many of the ideas in the book have been superseded by analysis of the Viking and later results, it provides good background on the evolution of scientific thinking about Mars. Many of the problems identified remain unresolved.

Squyres, Steve. *Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet*. New York: Hyperion, 2006. Written by the principal investigator for the Mars Exploration Rovers Spirit and Opportunity, this fascinating book provides a general audience with a behind-the-scenes look at how robotic missions to the planets are planned, funded, developed, and flown. A personal story of excitement, frustrations, a scientist's life during a mission, the satisfaction of overcoming difficulties, and the ongoing thrills of discovery.

Zubrin, Robert. *Entering Space: Creating a Spacefaring Civilization*. New York: Tarcher, 2000. The author displays a gung-ho attitude toward making humanity a truly spacefaring species by accepting the challenge of journeying to Mars sooner rather than later with contemporary technology and daring innovation. Speculates beyond travel to Mars.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Mars: Illusions and Conspiracies; Mars: Possible Life; Mars: Surface Experiments; Mars's Atmosphere; Mars's Craters; Mars's Polar Caps; Mars's Satellites; Mars's Volcanoes; Mars's Water; Meteoroids from the Moon and Mars; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits; Planetary Rotation; Planetary Satellites; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets.

Mars's Volcanoes

Categories: Mars; Planets and Planetology

The discovery of enormous volcanoes on Mars as a result of early images returned by the Mariner and Viking missions led to intensified study of Martian volcanic characteristics and activities. It is believed that this study will help scientists determine relationships between geological processes on Earth and Mars and develop a unified theory about the origin and evolution of the solar system.

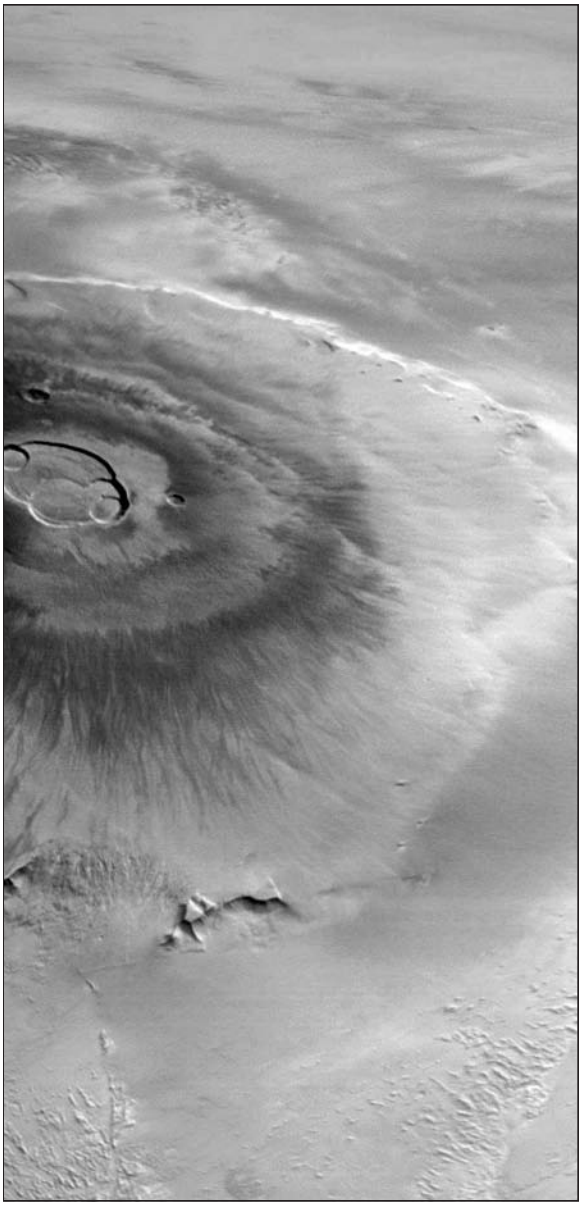
OVERVIEW

Volcanoes on Mars are generally larger than those on Earth. In fact, one should approach the planet with the understanding that all of its major features are gigantic, including its craters, plains, valleys, volcanoes, and polar caps. That is one of Mars's overall characteristics, and the enormous size of the Martian volcanoes is typical.

Observed Martian volcanoes seem to group themselves into two distinct regions: in the Tharsis region, atop the Tharsis Dome, and in the region known as Elysium, a large topographic region of crustal upheaval. Volcanoes in other areas tend to be smaller and older than those in these two regions, although the dispersion of Martian volcanoes is not broad. Sixteen principal volcanoes have been identified in these two prime areas, both of which are mostly in the planet's northern hemisphere. According to researchers, there are very few or no volcanoes in other regions of Mars, although the reason for this absence is not completely understood. Scientists speculate that the Martian crust is much thicker, overall, than Earth's crust and that it is much more difficult for magma to punch a hole through its surface. Also, the rigid Martian crust does not allow for tectonic plate movement, such as that which occurs on Earth.

Volcanoes on Mars are created in the same manner as terrestrial volcanoes, either by eruption through a central tunnel or by eruption through side vents in the volcano's walls. There is ample evidence that both processes have been at work on Mars in its long geologic past. In the first process, material from the planet's interior pushes up, overflows, and then cools rapidly, creating an upside-down cone with a hole in the top that eventually seals, plugging the eruption tunnel. In the second instance, magma oozes through breaks in the volcano's flanks, called vents. The major characteristics of Martian volcanoes, however, are very different from those of terrestrial volcanoes.

Researchers studying Mars have determined that there is no tectonic movement, or shifting of gigantic "plates" of crustal expanse, on Mars as there is on Earth. As a result, material from Mars's interior continues to erupt through the exact same tunnel over and over again, constantly building the volcano higher and wider. On Earth, on the other hand, a volcano is likely to be moved from its original spot, drifting away by the long-term motion of tectonic plates. This movement seals the magma tunnel deep inside the volcano, preventing further eruption.



Olympus Mons, the largest volcano in the solar system at more than half the size of Arizona, is also three times the height of Mount Everest, at 26 kilometers. The Mars Global Surveyor captured this image in 1998. (NASA/JPL/Malin Space Science Systems)

The second process, eruption of magma through side vents in the volcano's flanks, comes about by the interior materials bursting through the weakest points in the sides of the volcano. This kind of process occurs on Mars as on Earth, but the result is usually quite different. On Mars, lava flowing down the sides of the volcano spills beyond the volcano's base perimeter, creating vast lava plains that completely surround the volcano. Once again, because the volcano stays in one place, repeated eruptions through the side vents cause the lava plain to continue to grow broader and broader. On Earth, the buildup is considerably slower, if it occurs at all, because the volcano is moved away from its original eruption site by tectonic movement.

Volcanoes on Mars dwarf those on Earth. The largest of the Martian volcanoes are part of a grouping of about eight near the equator in the Tharsis Bulge, or Dome. One of the four biggest is Olympus Mons, a shield volcano judged to be about 27 kilometers high and 600 kilometers in diameter. The other three lie in a nearly straight line to the east of Olympus Mons. Beginning with the northernmost, their names are Ascraeus Mons, Pavonis Mons, and Arsia Mons. Although they are not as spectacular as Olympus Mons, their sizes are still quite impressive.

Martian calderas, large craters at volcanoes' summits caused by settling of the magma in their interiors, are similarly immense. The caldera of Olympus Mons, for example, is a highly complex collection of features measuring about 3 kilometers deep and 25 kilometers across, with its walls set at a slope of about 32°. The complex is the result of repeated caldera collapse after extrusions of magma settle and stop. Calderas of Martian volcanoes are unique, at least in the inner planets, in having many circular fractures surrounding the main caldera. In addition, the entire Olympus Mons structure is surrounded by a scarp, or cliff, which at some points is 6 kilometers high. In numerous places, there is ample evidence that lava has flowed over the sides of the scarp. Scarp lava overflows are often referred to as "flow drapes" and are quite common to Martian volcanoes; they extend the main volcano structure into the surrounding lava plain, often for many hundreds of kilometers. Flow drapes at Olympus Mons extend the volcano structure at least 1,000 kilometers outward over the rigid surrounding planetary crust. Flow drapes indicate that the scarp was formed before the time of the eruption of the magma which created the lava flows, and they help geologists to determine the age of the volcano and the extent of the extrusion of magma.

Terracing is a feature that is highly pronounced on the slopes of Martian volcanoes, created by the front lines of immense lava flows churning down a volcano's sides. Multiple terracing is often seen; it results from repeated eruptions, which create well-defined blankets of lava. Olympus Mons exhibits many such terracing features. It is difficult to imagine the enormous volume and extent of the rolling lava required to produce the many examples of terracing on this volcano.

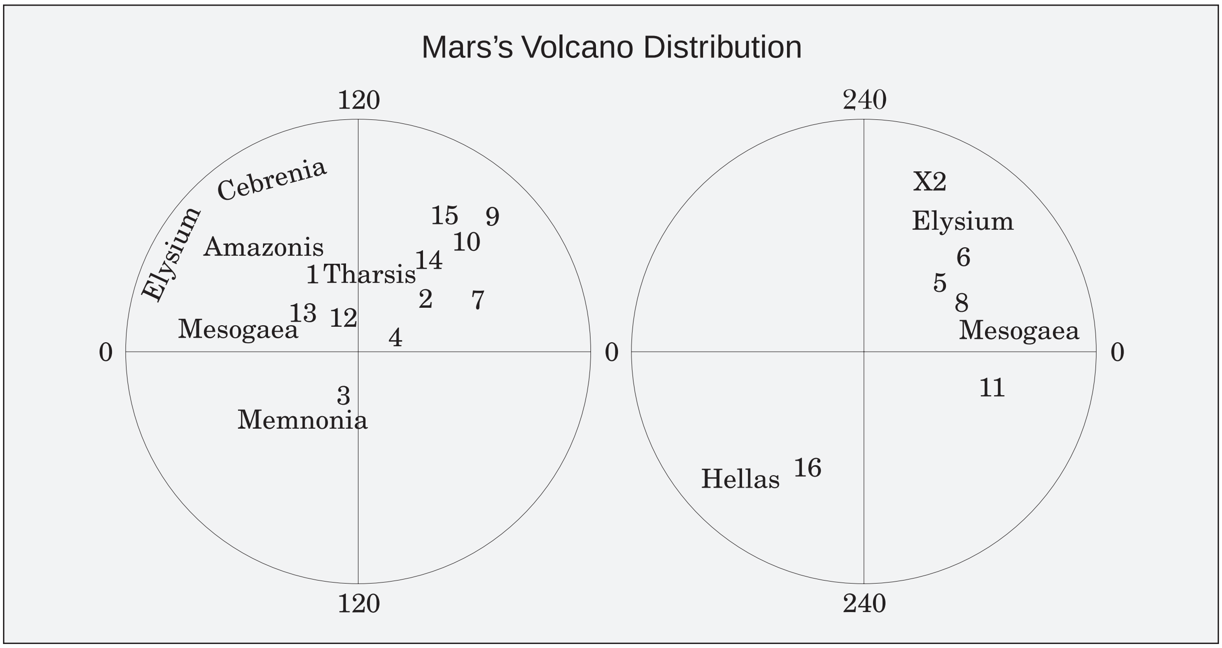
Arsia Mons rises some 16 kilometers above the Tharsis Dome with a caldera measuring 140 kilometers in diameter. The caldera is surrounded by concentric rings of hummocky topography, some of which are graben (long, linear depressions usually found between two parallel faults). The lava flow reaches at least 1,500 kilometers into the surrounding plain, often covering earlier features of the region. Images from the Mariner 9 mission, especially, disclose many lava-flow fronts and ridges where the rolling lava came to a stop. Repeated eruptions on Mars have thrown out truly copious amounts of lava, which, in turn, have formed far-reaching, relatively smooth plains adjacent to the volcano cones. These plains may extend hundreds of kilometers or, as in the case of Alba Patera, more than 1,700 kilometers out from the central cone. Although relatively flat, these plains most often show many lava-flow patterns, such as ridges and hummocks at the leading edges of flows.

The flow of lava down the flanks of a volcano can take many different forms. All these forms are present on Mars. The thickness of the lava, steepness of the slope, and presence or absence of barrier features all cause the downward spread of lava to take different forms and speeds, creating a variety of patterns. Fine-edged flows, flat-topped flows, and flow ridges can be seen in various volcano complexes. Ages of lava flows can be determined by the presence or absence of craters and other features on the volcanoes' flanks; numerous craters, for example, would indicate that the flows are relatively older and that the craters were formed after the lava flow, whereas lava that creeps into or flows over preexisting craters indicates that the flow came after the crater was created. Some ancient Martian volcanoes, such as Tyrrhena Patera, exhibit features that have been degraded by time, sometimes so much that it is impossible to determine the location of the primary volcano. Other features of some volcanoes suggest the downward flow of ash rather than of lava.

In sum, volcanoes on Mars are somewhat different from volcanoes on Earth. Martian volcanoes are significantly larger. The volcano cone is created differently, and the absence of tectonic movement allows a Martian volcano to stand over the same site for millions of years. Earth volcanoes in the process of eruption spew out lava in large amounts, yet Martian volcanoes eject lava in still larger amounts. When lava flows down the sides of a terrestrial volcano, it is affected by the landscape—streams, trees, and boulders which may deflect the flow pattern this way or that. On Mars, however, the flanks of volcanoes have no such barriers; instead, they encounter impact craters and the patterned buildup of previous lava flows. Finally, terrestrial volcanoes do not change size appreciably; those on Mars, however, keep growing bigger.

METHODS OF STUDY

Planetologists have enlisted a wide range of analytical techniques to study Martian volcanoes. The most obvious and by far the most productive have been the early images of the Martian surface gained through six highly successful uncrewed missions conducted by the National Aeronautics and Space Administration (NASA) between 1965 and 1976: Mariners 4, 6, 7, and 9 and Vikings 1 and 2. These flyby, orbiting, and surface-landing missions produced a huge amount of photographic material with which geologists have begun to piece together a profile of Mars's evolution based on the study of volcanoes. The study of volcanoes is especially enlightening because it allows a view of material that originated from deep inside the planet. The success of these early missions permitted the development of new questions in the study of Mars, and led directly to more advanced spacecraft such as Mars Pathfinder, Mars Global Surveyor, Mars Observer, Mars Reconnaissance Orbiter, Mars Express, and Mars Phoenix, many of which were committed in



large part to the search for water on Mars rather than volcanic processes.

By comparing Mars's topographic features with similar features on Earth, scientists can gain an enormous amount of information in a short time. By applying geometry to shadows and large-scale topography, they can determine the sizes and trace the short- and long-term development of evolutionary features. This information is critical in the case of volcanoes; lava-flow patterns, especially, are revealing about the structural evolution of volcanic sites and the surrounding terrain. Photographic evidence can be enlarged to a surprising degree, and often, extremely detailed images captured by remote-sensing technology can be studied simultaneously by scientists all over the world. Television cameras aboard the mission spacecraft gather and transmit digital image data, which are separated into shades of gray and then transmitted across space to receiving antennae on Earth. In special receiving stations, digital data are put together again and stored in computers. The data can be recalled from the computers at any time and can be manipulated by computer operators. Shadow areas can be lightened or darkened, color can be added and shaded, and mathematical computations can be applied to various features, all revealing the true nature of the topography.

The Viking landers, which descended to the surface of Mars, contained a multitude of scientific instruments that recorded data about wind direction and velocity, surface and atmospheric temperatures, sizes of windblown material, composition of Mars's soil, and even viscosity of the soil. There were two landers, each settling at a different location on Mars. An arm was extended from each spacecraft, and at the end of the arm was a specially designed scoop that dug a trench in the planet's surface. By analyzing the characteristics of the trench, scientists could deduce such information as the ability of the soil to cling together, whether or not other surface material slid into the trench, and the strength it took to dig the trench. Also, a series of three instruments aboard the landers received the collected soil, analyzed it chemically, and determined its composition. Data were collected in numerical formats and radioed back to Earth, allowing scientists to create a profile of the Martian environment. The profile was then applied to other regions on the planet, such as those where volcanoes existed. The picture of Mars that developed from these instrumented explorations is one of a cold, lifeless wasteland without the kind of environment that would allow humans to survive even for a short time.

Later landers provided circumstantial evidence for water and found samples of volcanic origin locally. The next quantum leap in understanding the volcanoes of Mars would be to send geologists to the Red Planet to survey volcanic regions and perform field geology exercises in situ.

CONTEXT

Properties of Martian volcanoes are important for scientists to study because they partially reveal the planet's present and past geology, providing clues to the planet's age, formation, and activity. Considerable study of volcanoes has disclosed that there is no tectonic plate activity associated with Mars, and that the Martian crust is far more rigid than Earth's. While there is a scarcity of volcanoes on Mars in comparison with the large number on Earth, they do help to establish theories about Mars's age and evolution. Volcanic ejecta can reveal something of the activities deep within Mars's crust and of the composition of material in the planet's mantle.

If scientists could decipher how Mars evolved into the planet it is today, they could compare it to Earth and the other planets and, perhaps, reach a greater understanding of Earth's past and future. One of the keys to unraveling Mars's mysteries is volcanic activity; it is a critical measure of what is or is not happening both on the planet's surface and deep inside its interior. Martian volcanoes are instrumental in the creation of other topographic features, such as extensive lava plains and scarps. Most volcanoes are located in a relatively confined area on Mars, and scientists want to know why volcanoes are not formed in other locations as well. Close study of Martian volcanoes could give scientists some clues about the interior properties of the planet and allow them to gain new insights into Mars's internal dynamics.

Commercially, Mars may harbor wealth in the form of natural resources. The Martian surface is characterized by planetwide iron deposits and iron oxide, from which Mars derives its pinkish-orange coloration. Other critical minerals and resources may exist there as well. It is important, therefore, to understand reasons for the present dispersion patterns of volcanoes in order to determine how volcanic actions affect mineral deposits and perhaps to predict future volcanic eruptions in mineral-rich areas.

Mars and, indeed, all planets must be viewed as large complex bodies with interacting physical systems and geological processes. Scientists, therefore, ask the following questions: What are the driving elements in the forces which sculpt, change, and characterize each planet, especially volcanic action? Is there a common thread among volcanic activities on all the planets and their satellites? How do these elements and forces behave on Earth? The study of Mars has enabled scientists to arrive at a new method known as Earth system science.

Thomas W. Becker

FURTHER READING

Barlow, Nadine. *Mars: An Introduction to Its Interior, Surface, and Atmosphere*. Cambridge, England: Cambridge University Press, 2008. An interdisciplinary text including contemporary data from the Mars Exploration Rovers and Mars Express. A great reference for planetary science students as well as the nonspecialist. Each chapter contains necessary background information.

Bell, Jim. *Postcards from Mars: The First Photographer on the Red Planet*. New York: Dutton Adult, 2006. An amazing number of high-quality black-and-white and color prints of Mars Exploration Rover images taken on the surface of Mars. The author was lead scientist for the rover's Pancam system; he shares the discovery process behind the photographs. Nontechnical.

Carr, Michael H. *The Surface of Mars*. Cambridge, England: Cambridge University Press, 2007. Heavily illustrated with comprehensive reference list. Author provides a complete description of the geological heating of Mars as understood based on results from Mars Global Surveyor, Mars Odyssey, Mars Reconnaissance Orbiter, Mars Express, Mars Pathfinder, and Mars Exploration Rovers.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic textbook on planetary science. Material about Mars covers all aspects of the Mars exploration. Takes a comparative planetology approach rather than presenting individual chapters on each planet in the solar system.

Hartmann, William K., and Odell Raper. *The New Mars: The Discoveries of Mariner 9*. NASA SP-337. Washington, D.C.: Government Printing Office, 1974. A classic companion volume to *Viking Orbiter Views of Mars* and an overview of the startling Mariner 9 mission, this book covers basic geology and planetology, comparing Martian and terrestrial features. Includes background data about Mars and its study, an explanation of the Mariner 9 mission sequence, and interpretations of the best of the mission's visual and scientific results. Graphs, charts, drawings, and photographs. Suitable for general audiences.

Kargel, Jeffrey S. *Mars: A Warmer, Wetter Planet*. New York: Springer Praxis, 2004. A member of Springer Praxis's excellent Space Exploration series. The book takes the reader on a search for Mars's water. The author provides a convincing case that the picture of a dry, waterless world portrayed initially by the early Mariner probes is not the Mars of today's understanding.

National Aeronautics and Space Administration. *Viking 1, Early Results*. NASA SP-408. Springfield, Va.: National Technical Information Service, 1976. The first scientific results of the Viking 1 Lander are chronicled in this booklet that represents a "first look" at the instrument and photographic evidence collected by a highly successful mission. Contains many mission photographs from both the lander and the orbiter. Charts, diagrams, maps, and graphs all contribute to a valuable reference work. For the advanced high school or college student.

Sheehan, William, and Stephen James O'Meara. *Mars: The Lure of the Red Planet*. Amherst, N.Y.: Prometheus Books, 2001. This book takes a different approach to the investigation of Mars, examining what it is about the Red Planet that is found so alluring. Also describes the great astronomers who advanced humanity's understanding of Mars from ancient times to the space age.

Squyres, Steve. *Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet*. New York: Hyperion, 2006. Written by the principal investigator for the Mars Exploration Rovers Spirit and Opportunity, this fascinating book provides a general audience with a behind-the-scenes look at how robotic missions to the planets are planned, funded, developed, and flown. A personal story of excitement, frustrations, a scientist's life during a mission, the satisfaction of overcoming difficulties, and the ongoing thrills of discovery.

Zubrin, Robert. *Entering Space: Creating a Spacefaring Civilization*. New York: Tarcher, 2000. The author displays a gung-ho attitude toward making humanity a truly spacefaring species by accepting the challenge of journeying to Mars sooner rather than later with contemporary technology and daring innovation. Speculates beyond travel to Mars.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Mars: Illusions and Conspiracies; Mars: Possible Life; Mars: Surface Experiments; Mars's Atmosphere; Mars's Craters; Mars's Polar Caps; Mars's Satellites; Mars's Valleys; Mars's Water; Meteoroids from the Moon and Mars; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits; Planetary Rotation; Planetary Satellites; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets.

Mars's Water

Categories: Mars; Planets and Planetology

Knowledge of how much water there once was on Mars and how much remains would shed light on the history of Mars and the solar system, the possible development of life in Mars's past, and ways of providing resources for future Mars colonists.

OVERVIEW

There is water on Mars. How much water and exactly where it is has become the center of many questions about the planet, including several mounted by the National Aeronautics and Space Administration (NASA). In 1971, the Mariner 9 spacecraft returned photographs of Mars that clearly showed that the surface of the planet had been extensively scarred sometime in its distant past, probably by liquid water flowing over it. Subsequent detailed investigations by the Viking probes in 1976 and the Mars Pathfinder rover in 1997 confirmed this evidence and added voluminous data to support the idea that liquid water had at one time existed on or close to the planet's surface.

Under present conditions, it is impossible for water to exist in liquid form on the planet's surface. The highest temperature recorded at the Viking landing sites on the warmest summer days was 244 kelvins, and the average temperature is much colder elsewhere on the planet. It fell to 150 kelvins at the site of the Viking 2 lander during Martian winter. Even if it were warm enough for liquid water to exist, it would quickly vaporize in the low atmospheric pressure of Mars. The atmospheric pressure at the Viking lander sites (the locations where the NASA probes made a soft landing on the planet's surface) was less than one one-hundredth of the Earth's atmospheric pressure at sea level. This pressure compares to that found at an altitude of about 35 kilometers above the Earth, nearly four times the altitude of Mount Everest.

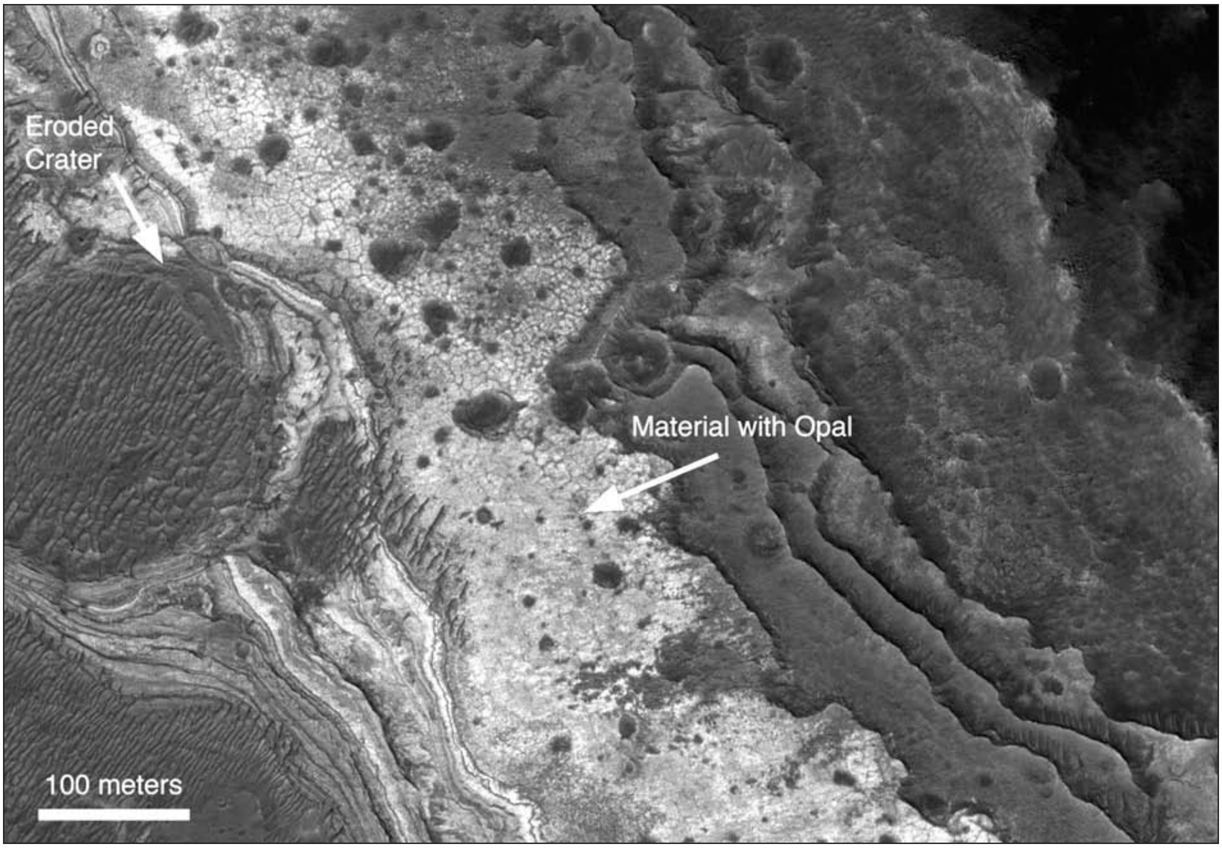
Water directly measured by the Viking orbiters was in a vapor state. There was also evidence that water exists in the solid form—as ice in subsurface permafrost, in the Martian polar caps, in the surface rocks and soil, in clouds and fog, and as frost. The amount of water vapor discovered in the Martian atmosphere was quite low compared to Earth standards. At the Viking lander sites, the atmosphere was 0.03 percent water vapor. This amount compares to the percentage of water vapor at an altitude of 9 or 10 kilometers on Earth. Mars is so dry that if all the water vapor in the atmosphere of the planet could be condensed into a solid block of water ice, it would measure only about 1.3 cubic kilometers—a tiny fraction of the water vapor in the Earth's atmosphere. The Viking probes measured the part of the Martian atmosphere with the greatest amount of atmospheric water vapor, which was the area nearest the planet's north pole.

Even though the Martian air is dry by Earth standards, because of the low atmospheric pressure it is near a saturated state; daily temperature variations cause the water vapor to condense in the form of ice fogs, clouds, and frost. The fog is quite tenuous; typical Martian fogs are about half a kilometer in depth, and if they were completely condensed into water, they would form a layer only a single micron thick. Frosts of water vapor, sometimes mixed with frozen carbon dioxide to form an ice called clathrate, form on the planet's surface seasonally.

The Martian poles are repositories of water ice and frozen carbon dioxide. The depth of water ice at the north and south poles varies. The north polar cap is estimated to be one meter to one kilometer deep. The south cap is less affected by the Martian summers, and its estimated thickness is 0.23 to 0.50 meter. It is primarily solid carbon dioxide. The amount of water ice in the caps is unknown. Polar water ice directly exposed to the Martian environment can undergo a deterioration by sunlight from solid directly to vapor by a process called thermal erosion.

Estimates show that Mars's original water, outgassed from the planet's interior by volcanism (as on the Earth), may have been sufficient to cover the planet with a layer some 46 meters deep. Most of it, however, was lost over the course of the planet's history in a process called molecular dissociation. High-energy ultraviolet sunlight split the water molecules into hydrogen and oxygen atoms, which were eventually lost to space. The equivalent of nearly 270,000 liters of liquid water is lost this way each Mar-

The Viking 2 lander took this photo of water ice on the rocks at Utopia Planitia in May, 1979. (NASA)



The Mars Reconnaissance Orbiter revealed hydrated rocks, similar to opal, both outside and within Mars's Valles Marineris canyons. It follows that water would have run on the surface two billion or more years ago. (NASA/JPL-Caltech/University of Arizona)

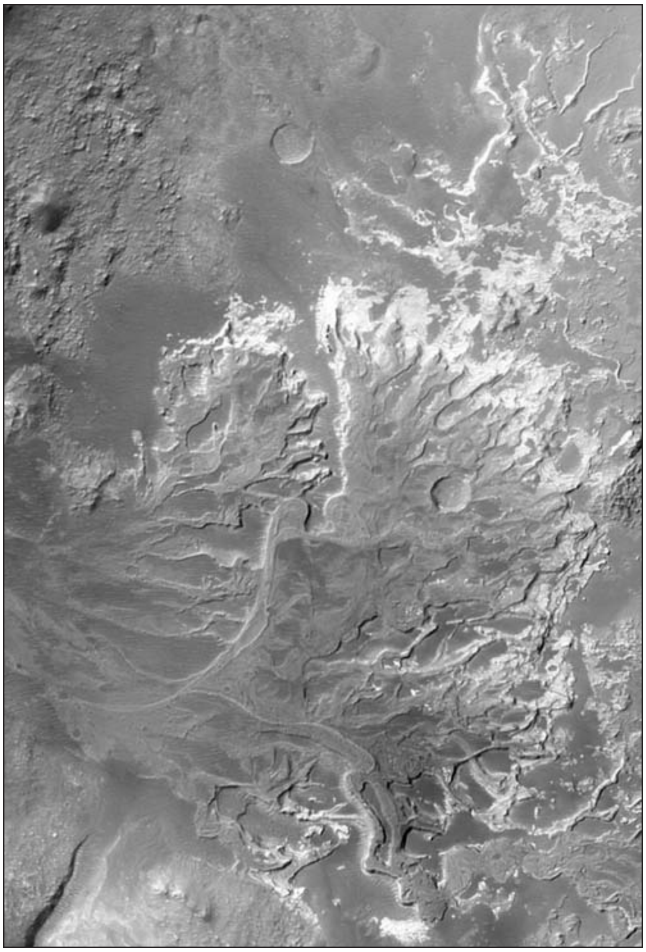
tian day. Orbital photographs show, however, that the surface is extensively scarred by what appear to be channels whose only Earthly analogue is caused by running water. Since liquid water requires an atmospheric pressure and temperature much greater that those which now exist on Mars, it is thought that the planet's conditions must once have been sufficiently different to allow water to flow. The Pathfinder mission supplied evidence of massive flooding.

As the orbital Viking probes circled Mars, they mapped seasonal variations of water vapor over the planet. It was discovered that as the spring and summer temperatures rose, water vapor levels increased also. Most scientists agree that the source of this water vapor was ground deposits of water. Water is encapsulated either in frozen aquifers or within the soil itself. Such frozen ground is called permafrost. Permafrost is located in the regolith of Mars; the depth of this upper layer of ice-laden soil is variable. Permafrost in the polar regions may be up to 8 kilometers deep; in the equatorial regions, it may extend to 3 kilometers.

Existence of large areas, perhaps planetwide areas, of permafrost is supported by orbital images that show four distinct geological formations: Meteoritic impacts display a kind of fluid ejecta, flowing away from the main crater, not seen on the Moon or the planet Mercury, that indicates that the energy of the meteoritic impact melted the permafrost. Another formation is called polygonal ground. These distinct polygonal shapes, observed from a high altitude, are caused by the repeated freezing and thawing of permafrost on or near the surface. Mass wasting (on Earth, often called landslides or mudslides) consists of distinctive downslope movements of soil, possibly caused by softening from water. Finally, a geological phenomenon called thermokarst was observed on Mars. It is caused by the underground melting of permafrost or underground ice formations, which lead to sinkholes or collapsed surface features.

In addition to permafrost, water is locked molecularly into the crystalline structure of the Martian soil and rocks. The Viking landers discovered that 0.1 to 1 percent of the surface materials consist of water of hydration. This tightly bound water can be released only by heating of the materials.

Many of the stream formations may have been caused by spring sapping during a planetary warming cycle. Spring sapping offers evidence of hidden, frozen deposits of ice that may thaw during warm cycles. The Mars Exploration Rovers Spirit and Opportunity provided evidence of sedimentary processes involving water. BB-sized spherules of gray hematite, referred to by the planetary scientists of the Mars Exploration Rover program fondly as "blueberries," were found by the Opportunity rover around the rim of Victoria crater. Formation of these gray hematite spherules occurs in the presence of water. Red hematite is just rust formation without the need of water, but the gray variety is found on Earth often in connection with hot springs where the oxidation of rust occurs in connection with water. However, it must be pointed out that gray hematite can be formed in connection with some types of volcanic activity, so the connection of the "blueberries" to water is not absolutely confirmed.



These fan-shaped, alluvial and delta-like deposits provide almost indisputable evidence of water flow on the ancient Martian surface, show that some sedimentary rocks were deposited in water, and strongly suggest that actual deltas existed where water met land. (NASA/JPL-Caltech/Malin Space Science Systems)

METHODS OF STUDY

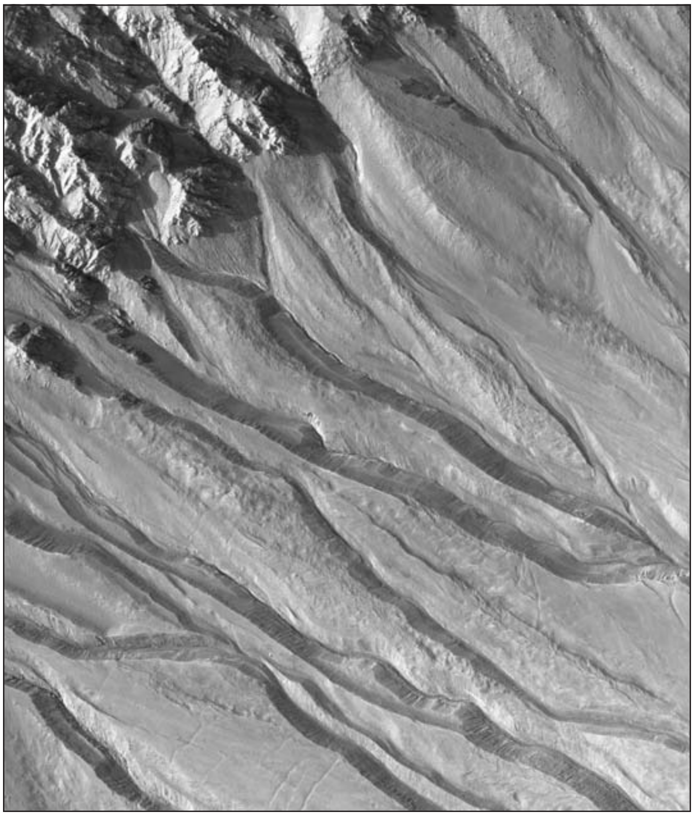
Most of the information regarding water on Mars came from eleven NASA spacecraft: Mariner 9, two Viking orbiters, two Viking landing craft, the Mars Pathfinder lander with its small independent rover Sojourner, Mars Global Surveyor, Mars Odyssey, Mars Reconnaissance Orbiter, and the Mars Exploration Rovers Spirit and Opportunity. The European Space Agency also began its expanding planetary exploration program by studying the Red Planet from orbit with its Mars Express spacecraft.

Direct evidence for Martian water came from orbiting instruments that measured atmospheric water vapor over seasonal periods and actually photographed the Martian polar caps, fogs, and cloud formations. Landing craft directly measured water vapor on the surface of Mars. They photographed seasonal frost deposits and clouds while measuring water of hydration by heating rock and soil samples. The instrument on the orbiter that measured the water vapor was called the Mars atmospheric water detector. It examined reflected solar radiation from the Martian surface at a spectral band of 1.4 microns.

The Viking lander instrument that analyzed the Martian soil for water was called the gas chromatograph mass spectrometer. Soil samples were taken from the surface of the planet by a robotic arm that directed the sample to a heating chamber. The soil was heated to 773 kelvins, and materials driven off by the heat were analyzed. It was discovered, after a series of samples had been analyzed, that between several tenths of a percent and several percent of the surface material was water. Some of the water was believed to be loosely absorbed on the surface, but it was likely that a significant fraction was accounted for by water of hydration.

Indirectly, scientists inferred much about Martian water repositories by comparing orbital photographs with high-altitude photographs of the Earth. Nearly all investigators were convinced that the only known mechanism that could form the clearly defined river- and streambeds was running water. This observation, coupled with simplified dating techniques of counting craters in streambeds to determine approximate ages, enabled speculation about cyclic Martian warming trends. These periodic warming trends could conceivably cause subsurface ice deposits to melt, and subsequent atmospheric pressure increases could allow the water to flow over the Martian landscape, cutting the stream formations in the soil.

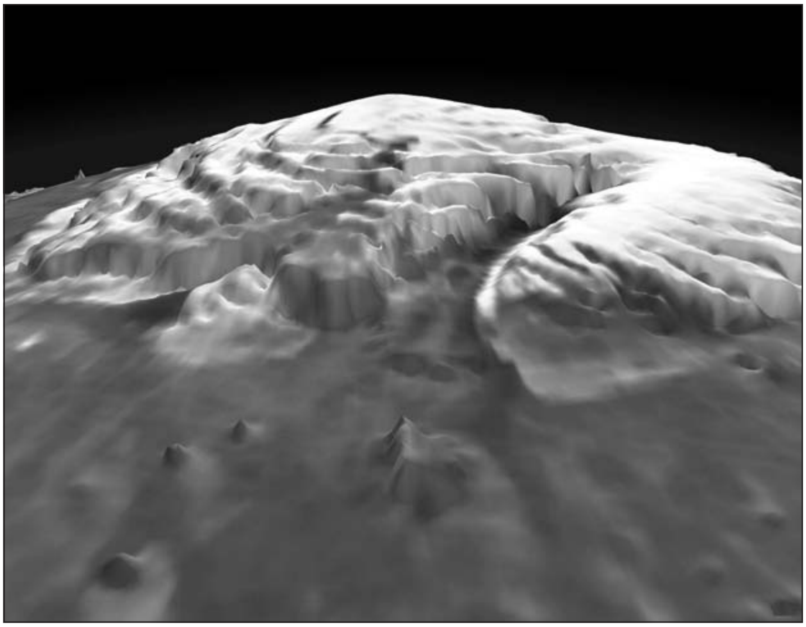
Examination of ejecta patterns all over the planet led to speculation that permafrost was a planetwide manifestation. Comparison of earthly geologic formations of polygonal ground, mass wasting, and thermokarst with their Martian counterparts provided evidence for the widespread nature and even depth of the Martian permafrost layer. Photographs of cloud and fog formation over the planet during subsequent orbits enabled the calculation of temperatures, saturation levels, and even the content of cloud and fog banks. After the lander data had verified the orbital photography, a highly accurate picture of Martian water deposits was formulated.



The southern highlands of Mars, in this image from the Mars Reconnaissance Orbiter, feature braided gullies typical of water channels. (NASA/Caltech/University of Arizona)

The Mars Polar Lander was sent to touch down in a polar region of Mars and search for direct evidence of subsurface water. Unfortunately, the spacecraft crashed and failed to transmit any data. The Mars Phoenix lander was designed to attempt the same thing. Launched in August, 2007, Mars Phoenix landed in the Red Planet's northern polar region on May 25, 2008, and began to dig in the soil to search for evidence of water. Mars Phoenix accomplished the first successful powered landing on Mars since the Viking touchdowns in 1976.

Mars Phoenix touched down in the northern polar region on May 25, 2008, at 68.2° north, 234° east, a location within what scientists had named the "Green Valley" of the Vastitas Borealis region. Locally, it was late spring at the time, but the surface temperature was still sufficiently cold that solid ice permafrost was strongly anticipated. After some difficulties with the lander's robotic arm and a few other critical systems, Mars Phoenix dug a trench in the Martian soil and exposed a white layer just below the surface. In time that white layer displayed sublimation, the phase change from solid directly to a gas. The rate at which the white layer sublimated strongly suggested that it must be water ice rather than dry ice (frozen carbon dioxide, which sublimates at an even greater rate at the local Martian temperature).



This three-dimensional image of Mars's north polar region has helped scientists determine the amount of water that may be in the ice cap. The image boasts a spatial resolution of 1 kilometer and a vertical resolution of between 5 and 30 meters. (NASA/JPL/GSFC)

Mars Phoenix was outfitted with a Thermal and Evolved Gas Analyzer (TEGA), essentially a combination of an oven and gas spectrometer. With TEGA, project scientists sought to detect water vapor released from heated samples of the white layer delivered to the hardware's oven chamber by actions of the robot arm and its scoop. Initial problems with clogging a TEGA sample inlet forced project scientists to be extremely careful in preparing for TEGA analysis. This delayed an unambiguous answer regarding whether or not the permafrost was water for several weeks, as Martian winter approached and threatened to shut down the Mars Phoenix lander.

It came as something of a relief when, on July 30, the Mars Phoenix lander's robotic arm delivered for the first time some viable subsurface material to an open chamber in the Thermal and Evolved Gas Analyzer. Several days of testing with the sticky Martian soil had produced a method whereby the arm's scoop could drop frozen permafrost into the vent leading down to a TEGA oven. The sample was heated in the oven and the evolved gases were analyzed. Just as the science team expected, the signature of water was confirmed. University of Arizona scientist William Boynton declared:

> We have water! We've seen evidence for this water ice before in observations by the Mars Odyssey orbiter and in disappearing chunks observed by Phoenix last month, but this is the first time Martian water has been touched and tasted.

In the wake of this discovery, NASA announced that funds would be forthcoming to extend the Mars Phoenix mission through at least September 30, a five-week extension.

Meanwhile, NASA's sequence of orbiters—Mars Global Surveyor, Mars Odyssey, and Mars Reconnaissance Orbiter (MRO)—trained a variety of instruments on the surface of the Red Planet. Their investigations were part of NASA's continuing program directive to search for water on Mars. MRO carried the largest cameras ever flown to another planet. The MRO data indicated the presence of underground water ice, and its photographs provided circumstantial evidence for recent changes in the surface where water may have played a role.

A pair of studies using MRO came to the conclusion that Mars once had water to the extent that large lakes and dynamic rivers existed for a long time in the distant past. In the July 17, 2008, issue of *Nature*, data were presented that showed that the Red Planet's ancient highlands, essentially 50 percent of the plant, contain clay minerals that can form only with water. Those clay features were later covered by volcanic lavas. However, the clay was uncovered across the surface by subsequent impact crater events. Features like the crater Jezero once confined a lake, and clay minerals were eroded down from the crater into a delta formation. The presence of these phyllosilicate minerals across the planet added fuel to the possibility that Mars once enjoyed wet environments that might have had the potential for development of primitive life.

CONTEXT

Mars Reconnaissance Orbiter studies in 2008 concluded that Mars in the past did have a wet environment as had been originally suspected after Mariner 9 images changed the scientific assessment of Mars. The newest studies suggested that Mars's ancient highlands contain clay minerals that form only with water. It appears that Mars had large lakes, vibrant rivers, and smaller wet regions that persisted for thousands and perhaps millions of years.

The question of what happened to Mars's water is of critical importance to the next generation of space explorers. It is also important to understanding how Earth's water reserves are balanced on a planetwide scale. Learning how nearly an entire planet's water resources, consisting of many trillions of liters, could simply vanish is critical not only to our insights into Mars's history but also to our understanding of Earth's future. Although the mechanisms of planetary water loss are understood, it is important to study Mars to learn exactly how and where the planet absorbed its remaining water resources. Scientists use techniques that could be employed to locate Earth's diminishing freshwater resources by space observations. Scientists may also learn how permafrost water deposits are linked to the contamination of water by soil salts and other impurities and how long-term climatic cycles lead to planetwide weather changes. If the Earth tilted only a few degrees, extreme global changes could be introduced that could have ramifications for the planet's long-term weather patterns.

Perhaps the most direct knowledge to be gained, however, is whether future colonists will be able to use what water there may be on Mars. Local water will be vital for the establishment of a Mars colony and will ultimately determine its size and usefulness. Water that is obtained from atmospheric distillation, the permafrost, the water of hydration, mining aquifers, or the polar caps will be used for a multitude of purposes, including drinking, agriculture, cooling equipment, washing and cleaning, and breaking down of molecular water into atomic hydrogen, for fuel, and oxygen, for breathing. Water on Mars may become one of the most significant aspects of "the desert planet."

Dennis Chamberland

FURTHER READING

Barlow, Nadine. *Mars: An Introduction to Its Interior, Surface, and Atmosphere*. Cambridge, England: Cambridge University Press, 2008. An interdisciplinary text including contemporary data from the Mars Exploration Rovers and Mars Express. A great reference for planetary science students and nonspecialists alike. Each chapter contains necessary background information.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. This beautifully illustrated and well-crafted book was intended to bring planetary discoveries to light in a single source. It covers the Martian water question in detail and extends it to a discussion of such ideas as life on Mars and the possibility of long-term Martian seasons. Written for the general reader.

Bell, Jim. *Postcards from Mars: The First Photographer on the Red Planet*. New York: Dutton Adult, 2006. An amazing number of high-quality black-and-white and color prints of Mars Exploration Rover images taken on the surface of Mars. The author was lead scientist for the rover's Pancam system. Shares the discovery process behind the photographs; not technical. Author states that his goal was to "share the beauty, desolation, grandeur, and alien strangeness" of Mars.

Carr, Michael H. *The Surface of Mars*. Cambridge, England: Cambridge University Press, 2007. Heavily illustrated with a comprehensive reference list. Author provides a complete description of the geological heating of Mars as understood based on results from Mars Global Surveyor, Mars Odyssey, Mars Reconnaissance Orbiter, Mars Express, Mars Pathfinder, and the Mars Exploration Rovers.

Collins, Michael. *Mission to Mars*. New York: Grove Weidenfeld, 1990. Apollo 11 astronaut Collins provides an astronaut's vision of a trip to Mars. Examines the problems to be overcome to make such a journey possible using space-shuttle-era technology.

Ezell, Edward Clinton, and Linda Newman Ezell. *On Mars: Exploration of the Red Planet, 1958-1978*. Washington, D.C.: Government Printing Office, 1984. This book is an official history of the Viking program, from its conception in 1958 to the culmination of the project some twenty years later. It is a detailed assessment of the program's political and technical history, but it also discusses details of the instruments that scanned for water, the Martian environment, and subsequent findings on the planet. Generally nontechnical and accessible to all readers.

Harland, David M. *Water and the Search for Life on Mars*. New York: Springer Praxis, 2005. A historical review of telescope and spacecraft observations of the Red Planet up through the Spirit and Opportunity rovers. Covers all aspects of Mars exploration, but focuses on the search for water, believed to be the most necessary ingredient for life.

Kargel, Jeffrey S. *Mars: A Warmer, Wetter Planet*. New York: Springer Praxis, 2004. A member of Springer Praxis's excellent Space Exploration series. The author provides a convincing case that the picture of a dry, waterless world portrayed initially by the early Mariner probes is not the Mars of today's understanding. The book takes the reader on a search for Mars's water.

Sheehan, William, and Stephen James O'Meara. *Mars: The Lure of the Red Planet*. Amherst, N.Y.: Prometheus Books, 2001. This book takes a different approach to the investigation of Mars, examining what it is about the Red Planet that is found so alluring. Also describes the great astronomers who advanced humanity's understanding of Mars from ancient times to the space age.

Squyres, Steve. *Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet*. New York: Hyperion, 2006. Written by the principal investigator for the Mars Exploration Rovers Spirit and Opportunity, this fascinating book provides a general audience with a behind-the-scenes look at how robotic missions to the planets are planned, funded, developed, and flown. A personal story of excitement, frustrations, a scientist's life during a mission, the satisfaction of overcoming difficulties, and the ongoing thrills of discovery.

Zubrin, Robert. *Entering Space: Creating a Spacefaring Civilization*. New York: Tarcher, 2000. The author displays a gung-ho attitude toward making humanity a truly spacefaring species by accepting the challenge of journeying to Mars sooner rather than later with contemporary technology and daring innovation. Speculates beyond travel to Mars.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Mars: Illusions and Conspiracies; Mars: Possible Life; Mars: Surface Experiments; Mars's Atmosphere; Mars's Craters; Mars's Polar Caps; Mars's Satellites; Mars's Valleys; Mars's Volcanoes; Meteoroids from the Moon and Mars; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits; Planetary Rotation; Planetary Satellites; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets.

Mercury

Categories: Mercury; Planets and Planetology

Mercury, the planet closest to the Sun, superficially resembles Earth's moon. Much that was known about this planet was obtained from experiments on board and photographic images returned by the uncrewed Mariner 10 probe, which completed three flybys of Mercury in the 1970's. A new round of investigations began early in the twenty-first century with MESSENGER mission.

OVERVIEW

Mercury completes one revolution about the Sun in only 87.97 days. Mercury's orbit has a mean distance from the Sun of only 0.387 astronomical unit (1 AU is the mean Earth-Sun distance), an eccentricity of 0.206, and an inclination of 7° with respect to the ecliptic plane. Mercury rotates about an axis with no obliquity and has a period of 58.65 days. The ratio of Mercury's rotational period to its revolution period is almost precisely two to three (2:3). Mercury's mass is 3.30×10^{23} kilograms, and its mean radius is 2,439 kilometers; therefore, Mercury's mean density is 5,420 kilograms per cubic meter. Mercury's most prevalent features are craters, scarps, and deformed terrain. Degradation of original craters has resulted from secondary impact and ballistic infilling, seismic activity resulting from impacts, lava flows, and isostatic readjustment.

Mercury's surface appears remarkably similar to that of Earth's moon, although Mercury's radius is about 50 percent larger than the Moon's. Both bodies are heavily pockmarked with impact craters. Closer examination, however, reveals many important differences between the surfaces of Mercury and the Moon. The Moon has greater color variations across its surface than does Mercury. Mercury's albedo, or reflectivity, is 0.12, a brightness similar to that of the lunar highlands seen on the Moon's Earth-facing side. Although there are 20 percent albedo contrasts across Mercury's surface, it lacks the dark maria and filled craters so prevalent on the Moon. On both worlds, younger craters are often higher in albedo and surrounded by prominent ejecta blankets and bright rays. On Mercury, most craters are less than 200 kilometers across. Many of the larger craters are double-ringed with flat floors that are usually shallower than their lunar counterparts. Central peaks are found in intermediate-sized craters, but larger circular features tend not to have central peaks. Other lunar features are absent or rather rare on Mercury. There appears to be no evidence of volcanic domes, cinder cones, or lava-flow fronts on Mercury. Rilles on the planet are usually straight rather than sinuous, are quite deep, and are as wide as 6 kilometers.

Mercury is surrounded by an extremely tenuous atmosphere of helium, argon, and neon. High daytime surface temperatures coupled with a low escape velocity lead to degassing, as the average thermal kinetic energy is sufficient to permit atmospheric escape in a relatively short period of geologic time. The average planetary surface temperature is 452 kelvins. However, the maximum dayside temperature is as much as 700 kelvins at closest approach to the Sun, and the minimum nightside temperature is 90 kelvins. Mercury exhibits the greatest equatorial temperature variation, more than 600 kelvins, of any planet in the solar system.

Mercury has a magnetic field only 1.6 percent as strong as Earth's. The origin of Mercury's magnetic field remains uncertain. A metallic core composed primarily of iron would be consistent with both the observed density and the magnetic field; however, Mercury's rotational speed could be too slow to generate currents in the core even if it is molten. Mercury's small magnetic field interacts with the solar wind. Mariner 10 recorded a moderately strong bow shock that traps energetic solar wind particles.

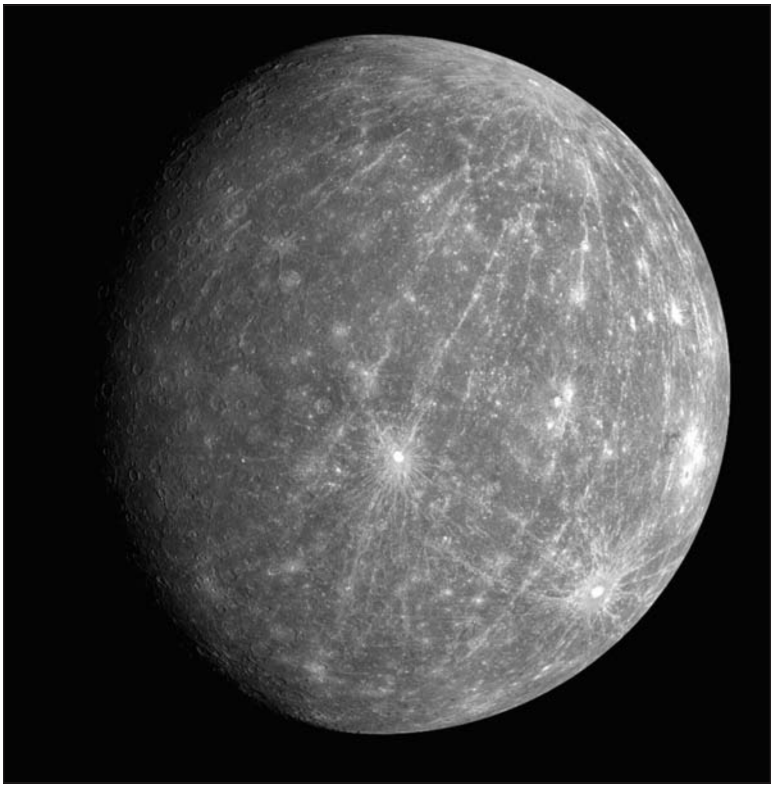
Although Mercury's surface resembles the Moon's, its mean density is closer to that of Earth. It is believed that, in proportion to its size, Mercury contains, in its core, double the amount of iron found in any other world in the solar system. High iron content would be consistent with the formation of Mercury by condensation of the solar nebula close to the Sun; heavy elements would have been more attracted to the early proto-Sun than would lighter elements. Existence of a magnetic field suggests that Mercury's iron has undergone differentiation and formed a hot, convecting core. Formation of the iron core would have generated heat in addition to the original radiogenic heat, causing expansion and melting of the entire mantle. This process would have had to occur before the era of intense cratering, because Mercury lacks surface expansion features and numerous preserved lava-flow formations.

According to one model of Mercury's interior, the planet's asthenosphere cooled quickly and thickened, possibly disappearing altogether. Mercury's lithosphere could extend down to the iron core, hundreds of kilometers below the surface. There is evidence to support such a model. After core formation, the planet would have cooled and contracted, resulting in a decrease in radius perhaps as large as one or two kilometers. Compression of the surface would then have caused thrust faults, the result of one rock unit slipping over another. Observed thrust faults on Mercury indicate a two-kilometer contraction.

Mercury's surface physiography can be classified into four major terrain types: heavily cratered terrain, smooth plains, intercrater plains, and hilly and lineated terrain. Intercrater plains are believed to be the oldest, predating the era of intense impact cratering. Smooth plains represent the youngest.

Smooth plains, located principally in the northern hemisphere near the large feature called Caloris Basin, are flat, lightly cratered surfaces akin to lunar maria. Craters in the smooth plains are typically sharp rimmed and only 10 kilometers across, at most. Some smooth plains fill the floors of large craters. Often, plains have sinuous ridges, an aspect shared by lunar maria. Regardless of location on the surface, Mercury's smooth plains have equal impact crater frequency, which indicates that smooth plain features were all formed at about the same time. Smooth plain features are believed to be volcanic in origin. There are too many smooth plains for them to have resulted from a single catastrophic impact or to be ejecta from the large Caloris Basin. Similarities between lunar maria and smooth plains suggest a common origin. Lunar maria are volcanic in nature, so smooth plains on Mercury are believed to be volcanic, also.

Intercrater plains, believed to be the oldest material on Mercury, form the largest physiographic feature on the planet. These plains have a greater crater density than smooth plains. Craters that pockmark rolling plains are typically less than 10 kilometers in diameter, and generally represent secondary rather than primary im-



Mercury, in this October, 2008, image sent back from MESSENGER, sporting massive impact craters and a series of rays from north to south. (NASA/Johns Hopkins University Applied Physics Laboratory/Carnegie Institution of Washington) pacts. Intercrater plains were formed by a variety of different events occurring over long periods. These plains are probably primordial crust that has been subjected to impact cratering, but their origin is by no means clear. The variety of intercrater plains suggests several alternative origins.

Heavily cratered terrain is reminiscent of the lunar highlands, being areas of many overlapping craters. Crater diameters vary between 30 and 200 kilometers. Ejecta deposits cannot be clearly identified with individual craters because of high degrees of overlap and disruption. This variety of terrain was formed as the era of intense bombardment was ending.

Hilly and lineated terrain is found directly opposite the Caloris Basin on Mercury and may have been formed by the Caloris impact event itself. This heavily deformed terrain, often referred to simply as Weird Terrain, covers about 250,000 square kilometers. It is made of hummocky hills 5 to 10 kilometers wide at the base and 0.1 to 1.8 kilometers high. Seismic energy from the Caloris impact apparently underwent antipodal focusing through the planet's core and broke or jumbled this region into hills and depressions.

The Caloris Basin is the largest single surface feature revealed by Mariner 10 photographs. More than 1,300 kilometers across, Caloris resembles the Moon's Imbrium Basin. It may represent an important event in Mercury's history, just as the Imbrium Basin does for the Moon. The basin is rimmed by mountains 30 to 50 kilometers wide and several kilometers high. Inside the basin are smooth plains scarred by small craters, ridges, and grooves indicative of lava flows modified by tectonic activity. Seeing the entirety of the Caloris Basin was a high-priority early objective of Mercury Surface Space Environment, Geochemistry, and Ranging (MESSENGER) flybys prior to beginning prolonged orbital studies.

Mercury Compared with Earth

Parameter	*Mercury*	*Earth*
Mass (10^{24} kg)	0.3302	5.9742
Volume (10^{10} km^3)	6.083	108.321
Equatorial radius (km)	2,439.72	6,378.1
Ellipticity (oblateness)	0.0000	0.00335
Mean density (kg/m^3)	5,427	5,515
Surface gravity (m/s^2)	3.70	9.80
Surface temperature (Celsius)	–170 to +390	–88 to +48
Satellites	0	1
Mean distance from Sun millions of km (miles)	58 (36)	150 (93)
Rotational period (hrs)	1,407.6	23.93
Orbital period (days)	88	365.25

Source: National Space Science Data Center, NASA/Goddard Space Flight Center.

METHODS OF STUDY

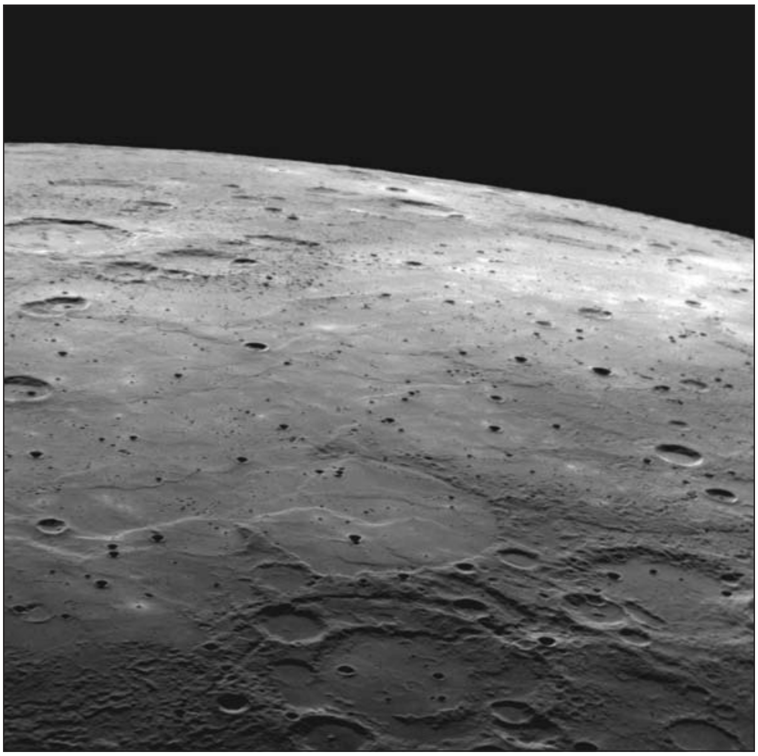
Mercury is a planet known to the ancients. However, Mercury reveals few of its secrets to visual observation from Earth. Because it is so close to the Sun, it is often hidden by solar glare and can be seen only briefly, visually or telescopically, at twilight or daybreak. The Hubble Space Telescope could not be used to obtain high-resolution images of Mercury's surface because of the tremendous brightness of the Sun, which would overpower and damage that orbiting observatory's sensitive instruments. Nevertheless, Mercury studies have advanced greatly since the first days of visual observations of this innermost planet in our solar system. Astronomers once incorrectly assumed that Mercury did not rotate as it revolved around the Sun. Few surface features were known before the Mariner 10 encounters. Indeed, for all intents and purposes, almost all that was known about Mercury prior to 2008 was obtained through scientific investigations performed by the Mariner 10 spacecraft on its three brief flybys in the mid-1970's. That probe was equipped with seven primary experiments; they involved high-resolution television imaging, infrared radiometry, radio wave propagation, extreme ultraviolet spectroscopy, magnetometry, plasma detection, and charged particle flux measurements.

Mercury's atmosphere was studied during a solar occultation using Mariner 10's ultraviolet experiment. The instrument measured the drop in the intensity of solar ultraviolet radiation as Mercury's disk and tenuous atmosphere obscured it. Data provided a profile of atmospheric concentration above the planet's surface. Other atmospheric data were gathered by monitoring radio waves emitted by Mariner 10 as it passed behind Mercury and then reemerged. The infrared radiometer, fixed to the spacecraft body on the sunlit side, had apertures which shielded the detectors from direct solar radiation. This experiment determined Venusian cloud temperatures as well as measuring surface temperatures on Mercury. Heat-loss data obtained as Mariner 10 crossed the planet's terminator, the line separating daylight from darkness, helped scientists infer information about the planet's surface composition. Surface brightness temperature was measured in a pair of spectral ranges, 34 to 55 micrometers and 7.5 to 14 micrometers, which represented temperatures of 80 to 340 kelvins and 200 to 700 kelvins, respectively.

Mariner 10 measured Mercury's magnetic field with a magnetometer package consisting of two three-axis sensors placed at different spots on a 6-meter-long boom. The use of two sensors provided the capability to isolate the spacecraft's own magnetic field from the weak field of the planet. Magnetic field measurements in interplanetary space were also made.

High-resolution images of any planetary surface can provide a wealth of information concerning that planet's past, its present geologic activity, and its surface composition. Mariner 10's television imaging system included two vidicon cameras attached to telescopes. The assembly was mounted on a scan platform that permitted the horizontal and vertical movements necessary for precise pointing. Cassegrain telescope systems were used in the imaging system. Powerful enough to resolve ordinary print at a distance of more than 400 meters, this system provided narrow-angle, high-resolution images. The television system also included an auxiliary optical system to obtain wide-angle, lower-resolution photography. This system was mounted on each of the television cameras. Experimenters were able to switch from narrow-angle to wide-angle imaging by moving the position of a mirror on the system's filter wheel. The vidicon cameras had 9.8-by-12.3-millimeter apertures and could make exposures of between 3 milliseconds and 12 seconds. Analog signals from the vidicon camera readout were digitized for transmission to receiving stations on Earth. An individual television image consisted of 700 vidicon scan lines, with each scan line consisting of 832 pixels.

The principal objectives of Mariner 10's television imaging program included collection of data useful in studying Mercury's planetary



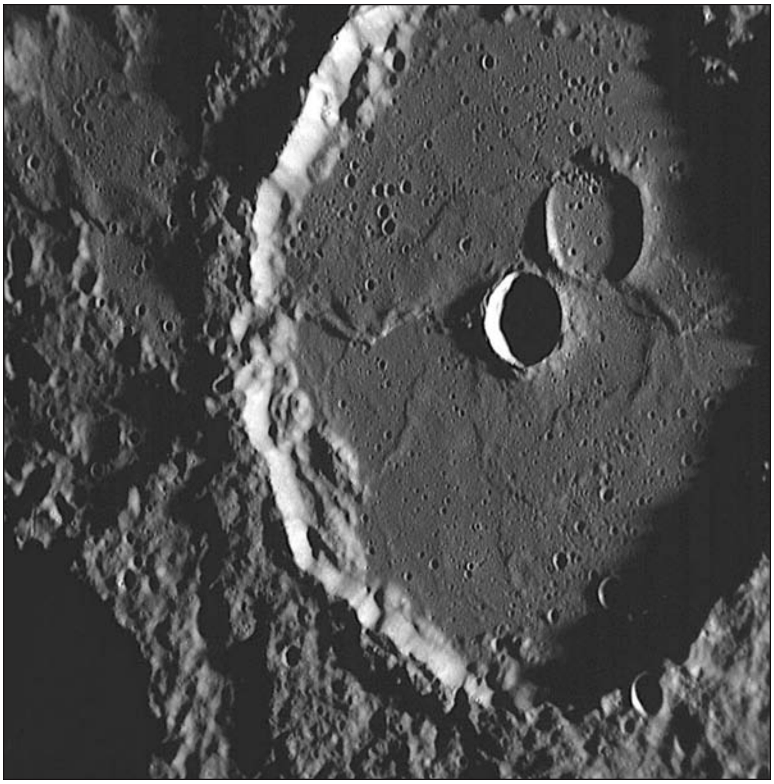
The MESSENGER 2008 second flyby produced this image of a previously unknown surface region, with both impact craters and tectonically smoothed terrain. (NASA/Johns Hopkins University Applied Physics Laboratory/Carnegie Institution of Washington) physiography, making a precise determination of Mercury's radius and rotation rate, evaluating Mercury's photometric properties, and categorizing the morphology of surface features. Television scans were made of the space surrounding Mercury in an attempt to locate unknown satellites, but none was found. This system was also used for studies of Venus and Comet Kohoutek before Mariner 10 even arrived near Mercury.

Data from Earth-based radar investigations of Mercury strongly suggest that at least part of the planet's core could presently be molten. Such a molten layer would have large implications for the production of the planet's global magnetic field and variations in Mercury's spin rate if the liquid core is decoupled from the solid mantle. In a 1992 issue of *Science*, Martin A. Slade et al. presented the results of two studies using the Arecibo radio telescope, the Very Large Array, and the Goldstone tracking antenna to send radio waves to Mercury at selected frequencies and detect the reflected signals. Essentially radar-astronomy exercises, the aim of both studies was to generate a radar reflectivity map of Mercury's surface at a resolution of about 15 kilometers. In the process, radar-bright returns that were highly depolarized were encountered near the planet's north and south poles. Data suggested the totally unexpected presence of ice on Mercury. Ice very effectively reflects radar at the gigahertz frequencies used in these studies and depolarizes those reflected radio waves greatly. Some of the radar-bright areas detected in these studies coincided with crater-sized spots. This provided evidence for the supposition that ice existed in crater areas that were permanently shadowed from solar radiation and therefore not heated tremendously, as was the rest of the planet when under daylight conditions. One of the bigger radar-bright areas was the large crater Chao Meng-Fu at Mercury's south pole. Planetary scientists supposed that the proposed ice came from either meteoritic bombardment or planetary outgassing (or both). Confirmation of this surprising result would have to await MESSENGER studies.

MESSENGER carried seven instruments that produced a great deal of data during the spacecraft's January 14 and October 6, 2008, encounters—information that would take longer to analyze than the time to the next and final flyby before eventual orbital insertion in 2011. The mission's objectives included photographing the as-yet-unseen 50 percent of the planet's surface, determining the composition and structure of Mercury's crust, understanding more about the planet's geological history, examining the planet's thin atmosphere, measuring the planet's quite active magnetosphere, searching for water at the poles, and providing data that would reveal the nature of Mercury's large core. This first attempt in thirty-three years to examine Mercury from close range was designed to help answer five separate major questions: What is the elemental and mineralogical composition of the surface? What does the surface look like at a resolution of better than hundreds of meters? What is the structure and temporal variation of Mercury's magnetic field? Does the planet's gravitational field exhibit any anomalies that might shed light on any uneven distribution of mass within Mercury? and What neutral particles and ions are found in Mercury's magnetosphere?

MESSENGER was outfitted with wide- and narrow-angle color and black-and-white imaging systems, a laser altimeter, a radio science experiment, and four multipurpose spectrometers. The spectrometers were capable of measuring spectra of gamma rays, neutrons, energetic particles and plasmas, and reflected light from Mercury's atmosphere and surface for compositional studies. The spacecraft's laser altimeter was designed to determine the elevation of the planet's surface features, as well as to look for wobble in the planet about its rotational axis. That sort of motion could help verify the existence of a suspected liquid layer in the core. The neutron spectrometer was designed to detect water ice at the polar regions. The laser altimeter was designed to measure the topography of permanently shadowed craters that might shelter water ice deposits. The ultraviolet spectrometer was designed to look for sulfur or hydroxyl deposits atop the water ice.

In May, 2008, researchers published results of laboratory modeling of Mercury's core that included a separated molten layer surrounding a



A close-up of the 106-kilometer-diameter Machaut Crater acquired on October 6, 2008, by MESSENGER. (NASA/Johns Hopkins University Applied Physics Laboratory/Carnegie Institution of Washington)

solid core. University of Illinois and Case Western Reserve University scientists hypothesized that deep within the planet an iron "snow" forms and moves down toward the solid core. Convection could be set up and create the planet's magnetic field. This experiment investigated the behavior of an iron-sulfur sample under tremendous pressure and heat. The iron-sulfur sample was set up to mimic the suspected core structure of Mercury. If formed, molten iron condenses to flake-like crystals, which would fall to the core. This heavy iron "snowfall" would result in lighter liquid sulfur rising, establishing convection currents. Observational data for MESSENGER will be able to determine if this laboratory model actually matches Mercury's internal structure.

In the meantime other scientists at Virginia Tech reported results of different simulations of conditions on Mercury. This work suggests that the shrinking of the planet's crust as Mercury cools over geological time should produce the thrust faults seen as scallop-edged cliffs and scarps on the planet. MESSENGER will also shed light on mantle convection, a process considerably different from that on Venus and Earth because of the thinness of Mercury's crust.

Details about the prolonged analysis of MESSENGER data collected during its first flyby surfaced in science journals in early July, 2008. The data confirmed that volcanic activity had played a tremendous role in the formation of Mercury's surface, especially during a period lasting from 3 to 4 billion years ago. MESSENGER provided evidence of volcanic vents along the margins of the Caloris Basin. Other evidence demonstrated that effusion had occurred. This process sees molten material from below the crust exude upward and outward across a planet's surface, sometimes forming features that resemble volcanic shields. Mercury had suffered lava floods that filled in fairly large craters almost to the wrinkled scarps outlining the craters. Some layers of lava were determined to be as deep as 2.7 kilometers.

CONTEXT

Mercury is the Roman name for the Greek god Hermes, patron of trade, travel, and thieves. Timocharis is considered to have registered the first recorded observation of Mercury, in 265 B.C.E. Very little more was learned about the planet until the invention of the telescope. Observation of the phases of Mercury was first reported in 1639 C.E. by Italian astronomer Giovanni Battista Zupus. Telescope technology improved, and evidence of surface features was found in the early 1800's, when astronomers Karl Ludwig Harding and Johann Schröter measured albedo variations.

It was not until the early 1960's that Mercury's rotation rate was precisely measured using radar observations. Then came the launch of Mariner 10, the final spacecraft in the historic Mariner series, on November 3, 1973, at 12:45 A.M. eastern time atop an Atlas-Centaur launch vehicle from Launch Complex 36B at Cape Canaveral. Photographs obtained during this flyby mission began the geologic analysis of Mercury. This spacecraft became the first to use gravity assists from large solar system bodies to redirect its trajectory to multiple photographic targets. It was recognized that the alignment of Earth, Venus, and Mercury was such that a single spacecraft could be launched between 1970 and 1973 from Earth toward Venus and then reach Mercury. Giuseppe Colombo of the Institute of Applied Mechanics in Padua, Italy, noted during an early 1970 Jet Propulsion Laboratory conference on the approved Mariner 10 mission that a 1973 launch opportunity existed in which the spacecraft could enter an orbit with a period nearly twice that of Mercury. That meant that a second Mercurian encounter was possible. Mariner 10 was indeed placed on a trajectory that permitted multiple encounters with Mercury, and this success demonstrated the feasibility of gravity-assist trajectories. The technique would prove tremendously valuable to the Voyager probes, which were sent to the outer solar system.

Shortly after Mariner 10's escape from Earth orbit, its planetary science experiments were activated to verify their operating condition. Mosaic photographs returned to Earth indicated that the spacecraft was in good condition to image a Moon-like world with high-quality camera systems. Mariner 10 came within 5,794 kilometers of Venus on February 5, 1974. During eight days of photography, the spacecraft returned 4,165 images of Venus and a wealth of data about the Venusian atmosphere. After another forty-five days of interplanetary cruising, the spacecraft reached the mission's principal target: the planet Mercury. Mariner 10 began taking photographs on March 23, 1974, reaching its closest approach, 5,790 kilometers, on March 29. The spacecraft then passed behind Mercury, to the nightside. More than two thousand photographs were obtained on this first encounter. Mariner 10's trajectory returned the spacecraft to Mercury on September 21, 1974; this time, it came as close as 50,000 kilometers. The probe completed a third encounter in March, 1975, before running out of fuel and entering a solar orbit.

The MESSENGER orbiter was designed to continue the scientific exploration of Mercury where Mariner 10 left off more than thirty years earlier. MESSENGER launched on August 3, 2004, and was injected into an interplanetary orbit that brought it back to Earth a year later for a gravity assist that would slow the spacecraft down to fall into the inner solar system. It was directed to encounter Venus in October, 2006, and again in June, 2007, for gravity assists that set it up for its first Mercury flyby. MESSENGER flew by Mercury near its equator on its first encounter, which took place January 14, 2008. Thus, little useful information about polar ice deposits was obtained. However, a highlight of the encounter was the capture of detailed images of the remainder of the Caloris Basin, not seen during Mariner 10's flybys. All spacecraft instruments functioned, signaling the potential for prolonged study once MESSENGER attained orbit. MESSENGER's flight path was refined by thruster firings so that it again encountered Mercury, flying by on October 6, 2008, collecting data and images and also setting itself up by a gravity assist in such a way that it would fly by Mercury one more time in 2009 before eventually entering orbit about the planet in 2011.

Even before entering orbit, as a result of the three flyby encounters MESSENGER was expected to give planetary scientists a nearly full initial map of Mercury's globe. To achieve orbit, MESSENGER's main propulsion system would fire to slow down by 860 meters per second. This fourteen-minute-long burn would consume 30 percent of the spacecraft's total fuel load. The first orbit would be adjusted until MESSENGER assumed an elliptical orbit ranging from 200 kilometers to 15,193 kilometers inclined 80° to Mercury's equator. In this nearly polar orbit, the spacecraft would orbit Mercury every twelve hours. Once in orbit about Mercury in March, 2011, MESSENGER's primary mission was to last four Mercurian years (the equivalent of one Earth year or two Mercurian solar days).

The European Space Agency (ESA) and Japanese Aerospace Exploration Agency (JAXA) also plan to investigate Mercury. Both space agencies intend to launch small probes in 2013. If successful, these will reach Mercury orbit in 2019. The ESA probe is named Bepi Colombo after the late Italian mathematician/engineer Guiseppi (Bepi) Colombo, father of the gravity assist. JAXA plans to send two more spacecraft to Mercury: one designed to conduct mapping operations, and another to investigate Mercury's magnetosphere. Although not designed to last that long, MESSENGER might exceed its mission life and be able to coordinate with these other spacecraft in Mercury studies.

David G. Fisher

FURTHER READING

Balogh, André, Leonid Ksanfomality, and Rudolf von Steiger, eds. *Mercury*. New York: Springer, 2008. This work provides background information and reviews changes in humanity's understanding about Mercury since the Mariner 10 flybys.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. Amply illustrated with color images, diagrams, and informative tables, this book is aimed at a popular audience but can also be useful to specialists. Contains an appendix with planetary data tables, a bibliography for each chapter, planetary maps, and an index.

Clark, Pamela. *Dynamic Planet: Mercury in the Context of Its Environment*. New York: Springer, 2007. Written by a NASA space scientist who edits the *Mercury Messenger* newsletter, this book covers the search for understanding of the solar system's closest planet to the Sun.

Davies, Merton E., Stephen E. Dwornik, Donald E. Gault, and Robert G. Strom. *Atlas of Mercury*. NASA SP-423. Washington, D.C.: National Aeronautics and Space Administration, Scientific and Technical Information Office, 1978. Provides an excellent contemporary description of Mariner 10 and its mission. Includes a full atlas of spacecraft photography of Mercury. An essential reference for the planetary science enthusiast or researcher; also accessible to general audiences.

Domingue, D. L., and C. T. Russell, eds. *The MESSENGER Mission to Mercury*. New York: Springer, 2008. A compilation of articles by experts in the MESSENGER mission to Mercury. Covers the flight, the science, the spacecraft, and engineering operations needed to conduct three flybys and insert MESSENGER into orbit around Mercury. For both the planetary science student and the astronomy enthusiast.

Dunne, James A., and Eric Burgess. *The Voyage of Mariner 10: Mission to Venus and Mercury*. NASA SP-424. Washington, D.C.: National Aeronautics and Space Administration, Scientific and Technical Information Office, 1978. This book offers an elegant description of the first spacecraft mission directed to the planet Mercury. Prepared immediately in the aftermath of the mission by the Jet Propulsion Laboratory, the text is complete with information on spacecraft operations and data returns. Photographs of Mercury abound. For general audiences.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that had investigated all planets with the exception of Pluto by the publication date. Takes an astrophysical approach to give our solar system a wider context as just one member of similar systems throughout the universe.

Moore, Patrick. *Moore on Mercury: The Planet and the Missions*. New York: Springer, 2006. An astronomical survey of humanity's knowledge of Mercury. Discusses what Mariner 10 observed and outlines the MESSENGER mission.

Spangenburg, Ray. *A Look at Mercury*. New York: Franklin Watts, 2003. Covers the search for understanding about the planet Mercury from antiquity through the Mariner 10 flybys. Previews the MESSENGER mission.

Strom, Robert G. *Mercury: The Elusive Planet*. Cambridge, England: Cambridge University Press, 1987. Dated, but covers all aspects of

Mercury as understood from Mariner 10 observations. Suitable for the general audience and younger readers.

Strom, Robert G., and Ann L. Sprague. *Exploring Mercury: The Iron Planet*. New York: Springer, 2004. According to a review by *Sky and Telescope*, this work is a "comprehensive text" that covers all that is known about Mercury. Contains a CD-ROM featuring Mariner 10 images and describes the anticipated MESSENGER mission.

Tanton, Linda Elkins. *The Sun, Mercury, and Venus*. New York: Chelsea House, 2006. A look at the innermost portion of the solar system and the star, our Sun, which plays such a prominent role in the evolution of both planets. For the general audience with an interest in science.

See also: Earth-Sun Relations; Europa; General Relativity; Gravity Measurement; Habitable Zones; Mars's Craters; Planetary Atmospheres; Planetary Magnetospheres; Planetary Orbits: Couplings and Resonances; Planetary Rotation; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Solar System: Element Distribution; Space-Time: Distortion by Gravity; Telescopes: Ground-Based; Terrestrial Planets.

Meteorites: Achondrites

Category: Small Bodies

Achondrites are a class of stony meteorites containing abundant silicate minerals that have formed as a result of igneous processes on small planetoids or asteroids. They closely resemble basaltic rocks found upon the Earth and Moon.

OVERVIEW

Meteorites are solid materials from outside Earth's orbit that have passed through the atmosphere and reached the surface. These objects—made of various combinations of rock and metal—are called meteoroids while still in orbit around the Sun. When a meteoroid encounters the Earth and enters the atmosphere, collisions with air molecules cause it to heat up and begin to vaporize. This produces a glowing trail of superheated air and hot vaporized material from the meteoroid that appears as a flash of light streaking across the sky. This phenomenon is properly called a meteor. (It also is commonly called a "falling star" or "shooting star," but such names are inaccurate since it is not a star at all.) It is estimated that the Earth is bombarded by hundreds of tons of meteoroidal material every day, but most meteoroids are no larger than small pebbles, and they vaporize completely in our atmosphere. If the meteoroid is large enough to survive its fiery plunge and land on the Earth's surface, then it is called a meteorite.

Meteorites are divided into three main groups based on the abundance of metallic and stony minerals they contain: the nickel-iron meteorites (often just called iron meteorites or irons), stony-iron meteorites (or stony irons), and stony meteorites (or stones). Stony meteorites are composed primarily of silicate and oxide minerals with minor amounts of metal. They can be subdivided into two subgroups known as the chondrites and achondrites.

Achondrites are not as common as chondrites; only about 1 in 10 stony meteorites is an achondrite. Achondrites get their name because they lack chondrules, mineral droplets that make up much of the material in chondrites. Achondrites have crystal textures similar to terrestrial igneous rocks, thus indicating that they formed when some larger parent body (perhaps the size of a small planet or large asteroid), melted, differentiated, and then cooled and solidified. Some achondrites have large mineral crystals that resulted from slow cooling as intrusive rocks inside the parent body. Others with smaller crystals were formed by more rapid cooling on or close to the surface of the parent body. Some resemble terrestrial lava flows that are riddled with bubble holes called vesicles, caused by gases that escaped. Some achondrites show evidence of collisions in space, resulting in a rock called an impact breccia that shows the effects of shock metamorphism. The pressure of impact causes the rocks to break apart and the minerals to shatter or deform, while the heat generated causes mineral and rock fragments either to melt slightly or to fuse together, depending upon its intensity.

Most achondrites are rich in one or more silicate minerals such as olivine (a magnesium-iron silicate), pyroxene (an iron-magnesium-calcium silicate), and plagioclase feldspar (a calcium-sodium-aluminum silicate) in varying proportions. Other minerals, such as spinel and chromite (iron-magnesium-aluminum-chromium oxides), are also found, as are small amounts of metal (less than 10 percent) in the form of iron and nickel alloys. In general, achondrites resemble a rock called basalt, which is a very common dark-colored igneous rock found on the Earth and Moon. With a few exceptions, achondrites are older and contain rare isotopes that make them very different from the rocks found on either the Earth or the Moon.

Achondrites can be subdivided into several types based on their texture and chemical composition. First and most abundant are the eucrites. They are similar in appearance to fine-grained terrestrial basalts, and they contain roughly equal amounts of calcium-rich silicate minerals such as plagioclase feldspar and pyroxene. In a hand specimen, a few eucrites exhibit a cumulate texture, which forms by the accumulation of coarse-grained crystals within a magma chamber. Still others possess a vesicular texture (containing many bubble holes or vesicles formed by escaping gases) and closely resemble terrestrial basaltic lava flows. Many eucrites also contain mixed fragments from other meteorite types and show the effects of shock metamorphism. The small variation in abundance of major chemical elements within all eucrites suggests their origin on the same parent body. They probably formed as extrusive and shallow intrusive igneous rocks later blasted out of the parent body by impacts. The age of eucrites has been dated using radioactive rubidium-strontium isotope techniques at 4.5-4.6 billion years, indicating crystallization very early in the history of the solar system.

Diogenites are achondrites that have a chemical composition similar to a terrestrial igneous rock called pyroxenite, which has an abundance of the mineral pyroxene. Texturally, diogenites have coarse-grained crystals that indicate slow, deep cooling below the surface. Based on laboratory melting experiments using actual achondrite samples, these crystals probably formed by cooling and crystallization from the same magma that also produced eucrites or by the more extensive melting of some eucrite source. Chemically, diogenites consist of metamorphosed accumulations of an iron-and-magnesium-rich but calcium-poor pyroxene known as a bronzite, along with minor amounts (less than 10 percent) of plagioclase feldspar crystals and some metallic iron. The bronzite crystals have become chemically homogeneous as a result of metamorphic heating. Like some eucrites, diogenites have been found shattered or mixed with pieces of other meteorites, resulting in a solid rock of angular broken fragments. Both diogenites and eucrites probably formed on the same parent body.

Howardites are a variety of achondrite that represents mixtures of many different meteorite types. They consist of crushed pieces from eucrites and diogenites, and they also contain about 2-3 percent by weight of pieces from chondritic stony meteorites. Texturally, howardites closely resemble the lunar soil. Under high magnification, a howardite's exterior surface is covered with small micrometeorite craters that contain impact-generated glasses, evidence of their formation by impacts on the surface of the parent body.

A likely parent body for these three types of achondrites seems to be the asteroid 4 Vesta. About 500 kilometers in diameter, it is the only large asteroid with a surface reflection spectrum like that of eucrites and diogenites. Several small Earth-approaching asteroids have similar reflection spectra, and they and the eucrites, diogenites, and howardites found on Earth probably were blasted off of Vesta by one or more large impacts.

Ureilites, another variety of achondrite, were named for the town of Novo Urei in Russia, where the first specimen was found in 1886. They consist of fairly large and abundant crystals of magnesium-rich olivine, some clinopyroxene, and a rare type of plagioclase feldspar set within smaller crystals composed of graphite, iron-rich metals, halite, sylvite, and troilite. In some specimens, the olivine crystals show a preferred orientation from crystal settling while molten; thus, ureilites exhibit variable textures. Most specimens have undergone intense high-pressure shock metamorphism that resulted in the formation of small diamonds from the graphite. Ureilites are the only achondrites that contain these tiny graphite and diamond crystals; the source of the carbon is unknown.

Aubrites, also known as the enstatite achondrites, are composed predominantly of a magnesium-rich pyroxene called protoenstatite and a rare type of plagioclase feldspar. Texturally, aubrites have large crystals, indicating slow cooling, but their origin and place of formation remain unexplained.

SNCs (pronounced “snicks”) are a small, unusual, and highly controversial group of related achondrites that include the shergottites, nakhlites, and chassignites. Shergottites are named for the town of Shergotty in the state of Bihar in India, where the first of these strange meteorites fell in 1865. Since that time, a few others like it have been found. As a group, the shergottites are similar to a terrestrial slow-cooled, coarse-grained igneous rock called diabase, which is rich in pyroxene and plagioclase feldspar. One of the feldspars found within the shergottites is maskelynite, a type of feldspar whose orderly atomic lattice structure has become disorganized from shock impact. Other minerals to be found are pyroxenes (calcium-rich augite and calcium-poor pigeonite), calcium- and sodium-rich plagioclase feldspars, oxidized iron in the form of magnetite, some olivine, and a rare water-bearing amphibole named kaersutite. Texturally, the shergottites are cumulates with elongated pyroxene crystals that have a preferred orientation, which is probably the result of flowage of newly formed crystals within the magma while still in a hot liquid state. Their geologic history records crystallization in a relatively Earth-like oxygen-rich environment and a period of intense shock metamorphism and high-intensity heating probably caused by impact, as indicated by numerous quickly cooled glass fragments. The radiometric age determinations on some of the shergottites’ minerals reveal a comparatively young age of 1.3 billion years. Trapped gas bubbles within some of the shergottite samples contain nitrogen and noble gases such as xenon, krypton, and argon, very similar in composition to the atmosphere of Mars.

The nakhlites are similar to terrestrial slow-cooled, coarse-grained igneous rocks known as gabbro. Mineralogically, these meteorites contain abundant augite (calcium-rich pyroxene) and smaller amounts of olivine, plagioclase feldspar, a few strange sulfide minerals, and metallic iron. Compared to the shergottites, all these minerals lack shock features, show no evidence of thermal metamorphism, but have a similar radiometric age of 1.4 billion years. Although their overall chemistry is different from that of the shergottites, nakhlites are believed to have been derived from either the same or a similar parent body.

Chassignites are named for Chassigny, France, where the first specimen was found in 1815. Several others have since been found in localities around the world. The few existing samples show that they are composed of abundant crystals of olivine with minor amounts of pyroxene, plagioclase feldspars, and kaersutite that show alteration by shock metamorphism. In a hand specimen, chassignites closely resemble terrestrial olivine-rich rocks called dunite.

Collectively, the SNCs contain cumulate crystals and a large percentage of volatile gases, which indicates formation upon a planet-sized body with a stronger gravitational field than that of the Moon. As a group, they have an average age of about 1.3-1.4 billion years, as determined by radioactive dating techniques. The chemical abundance and isotopic composition of gases trapped within small bubbles in these meteorites is nearly identical to the Martian atmosphere. The SNCs may have formed on Mars, from which they were blasted off into space by large impacts that hit Mars at the correct angle and speed to be eject material at speeds greater than 5 kilometers per second (Martian escape speed); these Martian rocks would eventually have encountered Earth and landed here as meteorites. Specimens from the surface of Mars will first need to be returned to Earth before it is possible to confirm that SNCs really originated there.

In 1974, a single strange meteorite named

Brachina was found in Australia. It is similar to the chassignites in mineralogy. Brachina has a fine-grained texture, contains 80 percent olivine and 10 percent plagioclase feldspar, lacks hydrous minerals, is unshocked, and is 4.5 billion years old (much older than any of the SNCs).

METHODS OF STUDY

Meteorites arrive daily on the Earth in sizes ranging from specks of dust to huge masses of several thousand kilograms. The vast majority fall into the ocean, never to be recovered, or into remote, uninhabited areas, to be discovered much later, if at all. Most achondrite meteorites probably are passed by unnoticed, because they closely resemble ordinary Earth rocks.

The best place to find meteorites of all types is Antarctica. In this remote, ice-covered continent, meteorites stand out starkly as black rocks against a white background of snow and ice. Because of the extreme cold and lack of liquid water, nearly all varieties of meteorites are found perfectly preserved. Meteorites are usually named for the closest town or post office in the vicinity where they are found; however, in the case of Antarctica, the name of the nearest mountain range, valley, or other topographic feature is used.

In the laboratory, the meteorite is weighed and measured, its density determined, and its physical appearance described. Thin sections are sliced from small chips of the meteorite for viewing under a petrographic microscope, where the behavior of light passing through the individual crystals of the specimen assists in identifying the minerals. The bulk chemical composition of the meteorite and a detailed analysis of its individual minerals can be made using an electron microprobe. When bombarded by an electron beam, the atoms within the specimen emit X rays. The atoms of each element emit X rays with characteristic energies, and the intensity at each X-ray energy indicates the abundance of the corresponding element. The overall texture of the meteorite and the distribution and abundance of each mineral present in it are used to place it in the classification system. Its bulk chemistry and elemental distribution help determine the processes that created it.

Extensive studies suggest that most achondrites probably came from water-free planetoids or asteroids. Based on meteorite melting experiments in the laboratory, the partial melting of a parent body with an overall chondritic composition could produce a eucrite. Other experiments using partially melted igneous rocks containing olivine and plagioclase feldspar produced magmas that, under the proper conditions, could form both diogenites and eucrites. Melting could easily have taken place in the low-pressure environment of space, provided that enough heat was generated by the decay of short-lived radioactive isotopes, such as aluminum 26, that once were abundant in these rocks.

Applying processes similar to those scientists believe formed the Earth, a parent body may be modeled that extensively melted and became partially separated into an upper layer of eucrite material atop a lower layer of diogenite material, surrounding a small, metallic iron-nickel core. The size of this parent body probably was no larger than a few hundred kilometers, so pressures on the interior core region would not have been more than 2 to 3 kilobars (2 thousand to 3 thousand times Earth atmospheric pressure at sea level). The mixing of eucrites and diogenites to form howardites probably occurred via meteorite impact, excavation, and lithification on or close to the surface of the parent body.

CONTEXT

Meteorites are samples of the building blocks of the planets. They have provided evidence for reactions in the solar nebula prior to the formation of the planets, processes occurring in planetlike bodies during their formation, and collisional impact events between solar-system objects. Continued study of the achondrites, along with other meteorite types, will provide more clues as to how the planets formed about 4.5 billion years ago.

Steven C. Okulewicz

FURTHER READING

Bevan, Alex, and John De Laeter. *Meteorites: A Journey Through Space and Time*. Washington, D.C.: Smithsonian Institution Press,

2002. A pictorial work that also deals with radioactive dating and geologic composition of meteorites. For general audiences.

Burke, John G. *Cosmic Debris: Meteorites in History*. Berkeley: University of California Press, 1986. Examines the role of meteorites in science history, from their origin as "thunderstones," their folklore and myths, and their curators and collectors to their role in current research. Many footnotes, a few photographs, and a detailed bibliography. Nontechnical and accessible to the general reader.

Dodd, Robert T. *Thunderstones and Shooting Stars: The Meaning of Meteorites*. Cambridge, Mass.: Harvard University Press, 1986. A thorough, clearly written review of information on all types of meteorites. Chapter 9 discusses achondrites and their parent bodies and how they relate to the origin of the planets.

Hutchison, Robert. *The Search for Our Beginning: An Enquiry, Based on Meteorite Research, into the Origin of Our Planet and Life*. New York: Oxford University Press, 1983. A clear and easy-to-read nontechnical book that describes the various types of meteorites and the information they provide about the solar system. Illustrated with many photographs of meteorites. Highly recommended.

McSween, Harry Y., Jr. *Meteorites and Their Parent Planets*. 2d ed. New York: Cambridge University Press, 1999. Gives an overview of the study of extraterrestrial debris. Good reference for scientists, students, and amateur astronomers.

Norton, O. Richard. *The Cambridge Encyclopedia of Meteorites*. New York: Cambridge University Press, 2002. Thoroughly discusses interior, external, and atomic compositions of meteorites. Valuable resource for scientists, students, and meteorite enthusiasts.

_______. *Rocks from Space: Meteorites and Meteorite Hunters*. 2d ed. Missoula, Mont.: Mountain Press, 1998. Introduces the reader to meteorites, asteroids, comets, and impact craters. This revised and updated edition includes dozens of new photographs.

Norton, O. Richard, and Lawrence Chitwood. *Field Guide to Meteors and Meteorites*. London: Springer, 2008. A guide designed to aid readers in observing meteors, as well as locating and analyzing meteorites.

Reynolds, Mike. *Falling Stars: A Guide to Meteors and Meteorites*. Mechanicsburg, Pa.: Stackpole Books, 2001. A guide for nonscientists on observing meteors and meteorites, including how to photograph them and how to record data.

Spangenburg, Ray, and Kit Moser. *Meteors, Meteorites, and Meteoroids*. Secaucus, N.J.: Franklin Watts, 2002. In addition to providing the basics about its topic, this volume also addresses meteoritic impacts with Earth, the process of watching and hunting meteorites, and what can be learned about the universe from them.

See also: Asteroids; Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Impact Cratering; Kuiper Belt; Lunar Craters; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Planetary Formation; Pluto and Charon; Solar System: Element Distribution; Solar System: Origins.

Meteorites: Carbonaceous Chondrites

Category: Small Bodies

The carbonaceous chondrite meteorites are the most primitive remnants of the primeval nebula from which the Sun, Earth, and all other bodies of the solar system originated. The hydrocarbon molecules present in these meteorites indicate the types of carbon-bearing molecules that most likely were present on the primitive Earth and may have been the building blocks of life.

OVERVIEW

Carbonaceous chondrites are a class of stony meteorites that are both chemically and physically primitive. They are chemically primitive in that, except for the elements hydrogen, carbon, oxygen, and the noble gases, the proportions of the elements in these meteorites are very similar to those observed in the Sun. They are physically primitive in that the carbonaceous chondrites escaped the thermal alteration (exposure to heat that causes changes in chemical composition and mineralogy) that affected almost all the meteorites in the other classes. Because of the primitive nature of the carbonaceous chondrites, they are thought to be the best samples currently available of the solar nebula out of which the Sun, Earth, and other solar-system objects formed. Thus, the composition of carbonaceous chondrites is generally taken as the starting point for models of the formation and subsequent evolution of the Earth.

The carbonaceous chondrites are relatively scarce, constituting only about 5 percent of all the meteorites recovered soon after their fall to Earth was actually observed. They are composed of millimeter-sized chondrules, individual grains whose mineralogy and texture indicate their crystallization from molten material, set in a matrix of finer-grained material. The carbonaceous chondrites are easily distinguished from all other meteorites by their dull black color, friability, generally low density, and almost total lack of nickel-iron grains, but not all carbonaceous chondrites are alike. Differences in composition, mineralogy, and texture allow the carbonaceous chondrites to be separated into several distinct types.

The most primitive type of carbonaceous chondrite, called the CI (or C1) type, is extremely rare, represented by only five meteorites. Of these, only Orgueil, a fall of about 127 kilograms, is large enough for extensive study. The others, Ivuna (0.7 kilogram), Alais (with little remaining of a 6-kilogram fall), Tonk (7.7 grams), and Revelstoke (only 1 gram), are all very small. The CI carbonaceous chondrites are different from all other chondrites, both carbonaceous and ordinary, in that they lack chondrules and consist almost entirely of low-temperature minerals, particularly clays.

Another type of carbonaceous chondrite, named CII (or C2 or CM), contains numerous organic compounds, including amino acids, that may have served as the basis for the development of life on Earth. At the very least they provide clues as to the types of organic material likely to have been present on the early Earth to serve as building blocks for life.

The CB carbonaceous chondrites exhibit a high oxidation state, with an abundance of volatile substances. The CO type is slightly less oxidized but contains metal and sulfides. The CV type closely resembles the CO type in mineral composition and oxidation state but contains large quantities of chondrules and whitish aggregates with a high calcium-aluminum content. Together, the CV and CO types are sometimes referred to as the CIII (or C3) type.

The carbonaceous chondrite types other than CI consist of a matrix, similar in chemical composition to bulk CI material, mixed with chondrules and aggregates of minerals that formed at high temperatures from a condensing gas and that exhibit a depletion in volatile elements from what is observed in the CI matrix. As the abundance of high-temperature material increases from about 1 percent in the CI type to about 60 percent in other types, the similarity of the bulk composition to that of the Sun decreases.

At one time, it was thought that the high-temperature material might be derived from the matrix by heating, which would eliminate the volatile material. Recent studies, however, show significant differences between the chemical and isotopic compositions of the high-temperature and low-temperature components of the carbonaceous chondrites, which make it impossible to derive one from the other by any simple process.

Much of the recent research on carbonaceous chondrite meteorites has focused on understanding the process by which these objects formed from the solar nebula, the gas and possibly dust that collapsed to form the Sun, Earth, and other solar-system objects. As the most primitive relics of the formation process currently available for laboratory analysis, the carbonaceous chondrites have been used to determine the chemical composition of the solar nebula, to establish the sequence and duration of events in the formation process, and to determine the temperatures characteristic of the process.

Major advances in the study of carbonaceous chondrites began in 1969. In February of that year, the Allende carbonaceous chondrite fell in northern Mexico, and about 2,000 kilograms of material were recovered for analysis. Later that same year, an even more primitive carbonaceous chondrite, the Murchison, fell in Australia. In addition, the return of lunar samples in 1969 spurred the development of research laboratories for the study of extraterrestrial materials. With the end of the Apollo lunar landing program, many of these laboratories shifted their emphasis to meteorite research and began to use highly sophisticated instruments perfected for lunar sample analysis to study meteorites.

Organic matter (chemical compounds of carbon, nitrogen, and oxygen) has been detected by spectroscopic methods in comets, on some asteroid surfaces, and on some planetary satellites. Study of the properties of this extraterrestrial organic matter would provide indications of the organic material likely to have been present on the Earth at the time life developed on this planet. The meteorites, particularly the carbonaceous chondrites, have been subjected to intensive examinations to determine whether they contain samples of this organic matter.

The search for organic matter in the carbonaceous chondrites was hampered for decades by terrestrial organic contamination of these meteorites from the time of their recovery until their analysis. By the time of the fall of the Murchison meteorite in 1969, researchers were aware of the contamination problem and efforts were made to preserve samples properly for organic analysis. In addition, several laboratories had recently developed procedures and instrumentation to search for organic material in returned lunar samples and to distinguish terrestrial contaminations from extraterrestrial organic material. The first analyses of the Murchison meteorite provided evidence for the presence of amino acids, which are the building blocks of proteins. Subsequent analysis of carbon in the organic matter indicated that the isotopic composition, the ratio of carbon 13 to carbon 12, was inconsistent with terrestrial contamination. It has now been demonstrated that several carbonaceous chondrites contain a varied suite of organic compounds. These same organic compounds have been duplicated in laboratory experiments by purely chemical processes and consequently are not evidence for life in space, but they are taken to indicate the types and variety of organic material likely to have been present on the Earth to serve as building blocks for life.

METHODS OF STUDY

Scientists have employed a variety of techniques and instruments to uncover the secrets locked in the carbonaceous chondrite meteorites. The chemical compositions, molecular abundances, mineralogies, isotopic ratios for individual elements, and present radioactivity have all been studied. Because of the small amount of carbonaceous chondrite material available for scientific study, especially of the rare CI type, many of the techniques employed to examine these meteorites have benefited greatly from the sophisticated instrumentation developed in support of the lunar sample analysis program.

The observation that the carbonaceous chondrites are primitive—that is, relatively unaffected by thermal processes—was established by detailed chemical analyses of individual mineral grains. The effect of prolonged heating is to cause the compositions of minerals of the same type to equilibrate, meaning that all grains of the same mineral from a single meteorite would have approximately the same composition if the meteorite were heated above the equilibration temperature. Such an effect is seen in most ordinary chondrites, but not in carbonaceous chondrites.

The compositions of small mineral grains are usually determined using an electron microprobe, an instrument that bombards the sample with an intense beam of electrons and detects the X rays emitted by the sample. When struck by an electron, each element emits X rays of specific energies. Thus, for example, the number of X rays emitted at the energy characteristic of the element iron gives the iron abundance in the sample. When the mineral grains in carbonaceous chondrite meteorites are examined by this technique, grains of olivine, the most easily altered of the major minerals in the matrix, exhibit a relatively wide range of compositions. In the Allende meteorite, for example, magnesium-rich olivine chondrules are found in direct contact with iron-rich matrix olivine. Such contacts eliminate the possibility that significant thermal events have occurred since the time at which the Allende meteorite was formed.

Observation of radioactive effects can also indicate the thermal history of these meteorites. The elements uranium and plutonium decay by nuclear fission, a process by which the nucleus splits into two fragments, each about one-half the mass of the original nucleus. These fragments fly apart with high energy, traveling a few thousandths of a centimeter before coming to rest. The host material can be damaged along the path of each fragment. This damage, called a fission track, is more easily attacked by reactive chemicals than is the surrounding mineral. After chemical etching, the fission tracks can be observed through a microscope. This damage, however, can be healed by heating. Thus, if fission tracks are revealed by chemical attack, the mineral has not been heated above the healing, or annealing, temperature after the fission event. The presence of tracks from uranium and plutonium fission in minerals from the carbonaceous chondrites indicates that they have not been heated above a few hundred degrees Celsius since their formation.

Similar radioactive decay processes can be used as clocks, providing a way to determine the ages of these meteorites. One such clock depends on the radioactive decay of rubidium 87, an isotope of the element rubidium, into strontium 87, one of the four stable isotopes of strontium. One-half of any initial sample of rubidium 87 decays to strontium 87 in 47 billion years. In any given sample, if the abundance of rubidium 87 as well as the amount of strontium 87 produced by radioactive decay could be measured, the elapsed time, or age, required for that amount of decay could be determined. In practice, application of this radioactive clock is complicated by a number of factors, including the migration of the strontium 87 from its decay site because of heating and the fact that not all the strontium 87 in the sample is from rubidium 87 decay. When appropriate corrections are made for these effects, however, the rubidium-strontium clock, as well as similar decay clocks using other pairs of elements, gives a consistent picture that the carbonaceous chondrites formed 4.55 billion years ago. Results for the oldest rocks on the Moon give essentially the same age. The extensive thermal activity in the early history of the Earth has apparently destroyed most or all evidence of the earliest rocks to form on this planet, but the observation that both the meteorites and the oldest lunar rocks have a common age suggests that the entire solar system, including the Earth, formed at that time.

Radioactive elements also provide clues to the duration of solar-system formation. Decay of aluminum 26, which is reduced to one-half of its starting abundance in only 720,000 years, produces magnesium 26. Magnesium has three stable isotopes, which are usually found in fixed ratios to one another, but the ratio of magnesium 26 to the other two isotopes of magnesium will increase when aluminum 26 decays. In some of the high-temperature aggregates from the Allende meteorite, significant enrichments in magnesium 26 were found by mass spectrometry. Detailed examination of the minerals containing these enrichments showed that the size of the magnesium 26 enrichment increased in proportion to the aluminum concentration in that mineral. This suggested that radioactive aluminum 26 was incorporated into the mineral and subsequently decayed to magnesium 26. For this to be true, however, the high-temperature aggregates would have to have formed within a few million years of the isolation of the solar nebula, or most of the aluminum 26 would already have decayed. Thus, the high-temperature aggregates in Allende and some other carbonaceous chondrites provide evidence that mineral grains condensed very early in the solar-system formation process.

CONTEXT

As a group, the carbonaceous chondrites have provided a wealth of information on many different aspects of planetary formation and the development of biological processes that may have occurred in the early solar system. They exhibit a variety of conditions of formation that range from a high-temperature, low-pressure, volatile-poor environment to one that was at a lower temperature and volatile-rich. They appear to have experienced only minor alteration since their formation from the collapsing gaseous nebula that became our solar system, and thus they preserve a record of that early era of solar-system history. The chemical composition of the least altered of these meteorites is almost identical to the composition of the Sun, except for a few gaseous elements. Thus, the carbonaceous chondrite composition is taken to indicate the bulk composition of the Earth, which cannot be measured directly since the Earth's interior is inaccessible.

Radioactive clocks in the carbonaceous chondrites indicate that they formed 4.55 billion years ago. The consistency of this age with the age of the oldest rocks brought back from the Moon by the Apollo lunar landings is taken to indicate that the entire solar system, including the Earth, formed at that time. Isotopic relics of other radioactive elements, now extinct, demonstrate that some minerals in the carbonaceous chondrites formed within as little as a few million years of the isolation of the solar nebula from addition of new galactic radioactive isotopes.

The carbonaceous chondrites also contain organic molecules, including amino acids, which are the building blocks of proteins. Although there is no evidence of biological activity on the parent body of the carbonaceous chondrites, these or similar organic molecules are likely to have been available on Earth to serve as building blocks for the development of life.

George J. Flynn

FURTHER READING

Dodd, Robert T. *Meteorites: A Petrologic-Chemical Synthesis*. London: Cambridge University Press, 1981. This is a well-illustrated summary of the mineralogical and chemical analyses of all types of meteorites. Chapter 2 focuses on the chondritic meteorites and their relation to one another. Chapter 3 describes the properties of the carbonaceous chondrites, their relationship to the solar nebula, and the possibility that they contain presolar grains. Text is suitable for college-level readers who have a minimal background in Earth science.

_______. *Thunderstones and Shooting Stars: The Meaning of Meteorites*. Cambridge, Mass.: Harvard University Press, 1986. This reference book explains why there is a scientific interest in meteorites and summarizes what is known about them. It includes chapters on the chondritic meteorites and on the parent bodies of these meteorites. While written as a college-level text, it provides detailed explanations of the phenomena and techniques of analysis without requiring the reader to have an Earth science background.

Erickson, Jon. *Asteroids, Comets, and Meteorites: Cosmic Invaders of the Earth*. New York: Facts On File, 2003. Part of The Living Earth paperback series. Discusses the threats that these objects can pose to Earth and the devastating effects that major impacts have had, and still pose, to life on Earth. Also provides a general description of these bodies separate from their threat of impacting the planet. For the general reader.

Hutchison, Robert. *The Search for Our Beginning: An Enquiry, Based on Meteorite Research, into the Origin of Our Planet and Life*. Oxford, England: Oxford University Press, 1983. Summarizes the present state of scientific knowledge about the Earth, the Moon, and the inner planets and describes how the knowledge gained from the study of meteorites has shaped theories of the origin and evolution of the inner solar system. Illustrated; intended for general readers.

Kerridge, John F., and Mildred S. Matthews, eds. *Meteorites and the Early Solar System*. Tucson: University of Arizona Press, 1988. A collection of articles by sixty-nine contributing authors describing the state of meteorite research as it relates to early solar-system processes. Well illustrated; includes a detailed review and comprehensive bibliography of the major topics in meteorite research accessible to general-level readers. Aimed at graduate students.

Mason, Brian. *Meteorites*. New York: John Wiley & Sons, 1962. This book emphasizes the chemical and mineralogical measurements made on the meteorites and describes the relationship between meteorites and other objects in the solar system. It provides a state-by-state listing of all the meteorites collected in the United States up to the date of the book's writing, including the exact location, year of recovery, weight, and type.

Nagy, B. *Carbonaceous Meteorites*. New York: Elsevier, 1975. This book describes the carbonaceous chondrite meteorites, with particular emphasis on the carbon-rich phases, and discusses the long effort to identify organic compounds in meteorites. Intended as a college text, this book is suitable for readers with a high school science background.

Wasson, John T. *Meteorites: Classification and Properties*. New York: Springer, 1974. A college-level introduction to meteorite research and classification by type. The emphasis is on interpretation of the chemical and mineralogical data. Well illustrated; includes a tabulation of all meteorites known by the early 1970's, with the type of each meteorite indicated. Appendix C lists each carbonaceous chondrite known at the time of publication.

_______. *Meteorites: Their Record of Early Solar-System History*. New York: W. H. Freeman, 1985. This college-level text is less technical than Wasson's 1974 book. This well-illustrated volume describes the formation processes for the different meteorite types and attempts to link the different meteorite groups with their appropriate parent bodies. Chapter 7 provides an extensive discussion of the chondritic meteorites, including the carbonaceous chondrites.

See also: Asteroids; Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Impact Cratering; Kuiper Belt; Lunar Craters; Meteorites: Achondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Planetary Formation; Pluto and Charon; Solar System: Element Distribution; Solar System: Origins.

Meteorites: Chondrites

Category: Small Bodies

Chondrites are a class of stony meteorites that contain chondrules, small mineral droplets that make up a prominent part of their material. Chondrites are the most common type of meteorite.

OVERVIEW

Meteorites are divided into three main groups based on the abundance of metallic and stony minerals they contain: the nickel-iron meteorites (often just called iron meteorites or irons), stony-iron meteorites (or stony irons), and stony meteorites (or stones). Stony meteorites (or stones) are the most abundant of the three groups. They are composed mostly of the silicate minerals olivine, pyroxene, and plagioclase feldspar. Metallic nickel-iron grains occur in varying small amounts and are accompanied by an iron-sulfide mineral called troilite, which is very rare on Earth. Stony meteorites have the greatest variety in composition, color, and structure.

One particular structural feature called chondrules divides the stony meteorites into two main subgroups: the chondrites, those with chondrules, and achondrites, those without. Chondrites are much more common than achondrites; about nine of ten stony meteorites are chondrites. The words "chondrule" and "chondrite" are derived from the Greek word *chondros*, meaning grain. Chondrules are small, rounded particles generally made of high-temperature silicate minerals whose texture indicates rapid crystallization from molten material. These chondrules range in size from less than a millimeter to just under a centimeter. They can be either whole or partial and are embedded within a matrix of fine-grained opaque minerals and glass. The most common types of chondrules are the barred olivine, the excentroradial "feathery" pyroxene, and the porphyritic olivine and pyroxene varieties. These mineral textures, along with the interstitial glass, are indicative of rapid cooling from a high-temperature liquid.

Mineralogically, most chondrites typically are composed of about 45 percent olivine, 25 percent pyroxene, 10 percent plagioclase feldspar, 5 percent troilite, and 2 to 15 percent nickel-iron alloy minerals (kamacite and taenite). Chondrites that tend to be poorer in olivine have a correspondingly higher metal content, and the reverse is also true. This relationship gives rise to a classification scheme in which the three main groups of chondrites are defined: the H type (high iron content), the L type (low iron content), and the LL type (very low iron content). Visually, specimens may be roughly classified into one of these groups by the amount of metal observed on a cut surface. Texture also can play a part in identifying specific types, as the LL type tends to be more brecciated (composed of rock fragments) than the H and L types. These three groups make up the vast majority of chondrites, but other, rarer groups are also recognized. The E type (or enstatite) chondrites are named for the predominance of the mineral enstatite (an iron-free magnesium-rich pyroxene); iron, constituting about 15 to 25 percent of this type's composition, commonly occurs in the metallic state. The principal differences in the mineralogy of these chondrite types can be attributed to different oxidation states and thus reflect specific conditions of formation. The H, L, and LL chondrites are indicative of a relatively high oxidation state. The E chondrites, on the other hand, are indicative of more reducing conditions and offer an interesting insight into variations in meteorite formation processes.

Another rare group, the carbonaceous chondrites (C type), derives its name from the carbon found in their bulk chemistry, which far exceeds the normal trace amounts found in the ordinary chondrites. Most chondrites (the ordinary chondrites) contain little or no carbon, but a few (the C type, or carbonaceous chondrites) contain a suite of carbon compounds. The carbonaceous chondrites exhibit a wide variation in their chemistries and conditions of formation. These differences divide them into several distinct types. The most primitive type of carbonaceous chondrite, called the CI (or C1) type, is different from all other chondrites, both carbonaceous and ordinary, in that carbonaceous chondrites lack chondrules and consist almost entirely of low-temperature minerals, particularly clays. The CII (or C2 or CM) type contains numerous organic compounds, including amino acids, that may have served as the basis for the development of life on Earth. The CB carbonaceous chondrites exhibit a high oxidation state, with an abundance of volatile substances. The CO type is slightly less oxidized but contains metal and sulfides. The CV type closely resembles the CO type in mineral composition and oxidation state but contains large quantities of chondrules and whitish aggregates with a high calcium-aluminum content. Together, the CV and CO types are sometimes referred to as the CIII (or C3) type. The carbonaceous chondrites exhibit a variety of conditions of formation that range from a high-temperature, low-pressure, volatile-poor environment to one that is at a lower temperature and volatile-rich. As a group, the carbonaceous chondrites have provided a wealth of information on many different aspects of planetary formation and the development of biological processes that may have occurred in the early solar system.

Chondrites show varying degrees of similarity to Earth rocks, yet they are distinct from any Earth rock. Their basic mineralogy reveals that most chondrites were originally formed under high-temperature conditions from molten material like igneous rocks but were later broken up and reassembled, like sedimentary rocks. Once incorporated into a new mass, the chondrites were subject to variations in temperatures and pressures, thus becoming metamorphosed. It is evident that they have had a rather complex evolutionary history and can therefore provide interesting evidence for the interpretation of early solar-system history.

The study of chondritic meteorites therefore raises important questions based on variations in their chemical compositions and textural features. How did they condense from the solar nebula, and what type of parent body produced all these variations? To try to provide some answers, a hypothetical parent body between 200 and 300 kilometers across is proposed. This body would have been chondritic in nature and subjected to partial melting, presumably because of the short-lived radioisotope aluminum 26. Later bombardment by smaller bodies would produce localized melting and the brecciation that is common to most chondrites. This model is highly speculative but does offer a reasonable explanation for many of the features of the different types. In addition, some of the chondrites retain some of their pre-parent-body characteristics and offer evidence of the conditions that existed before the accretion of asteroid-sized bodies. Thus the chondritic meteorites play an important part in the understanding of planetary formation.

METHODS OF STUDY

Chondritic meteorites are studied by many different analytical methods. The first is the determination of bulk chemical content. This can be achieved through basic wet chemical techniques or by the use of more sophisticated methods such as neutron activation analysis and X-ray fluorescence. Where individual minerals are large enough, X-ray diffraction can be employed for a positive identification of the particular mineral phase. However, the mineral grains in most chondrites usually are too small and require an alternative approach using an electron microprobe. This instrument utilizes a microscope to locate tiny individual mineral grains, and then electron bombardment of the grain causes the atoms of each element in it to emit X rays with characteristic energies, making it possible to determine very precise mineral phases at the microscopic level.

Once the meteorite's mineralogy is known, it becomes possible to employ radioisotope dating techniques to learn the age at which the minerals crystallized. This can be achieved by use of potassium-to-argon (K-Ar), rubidium-to-strontium (Rb-Sr), uranium-to-lead (U-Pb), and thorium-to-lead (Th-Pb) radioactive decay rates. It is from such data that some chondritic meteorites have been established as the oldest known solid materials in the solar system.

To determine what has physically happened to the chondrite over the eons requires optical analysis with a petrographic microscope. This instrument offers a magnified view of the texture of the meteorite and the minerals it contains, which in turn provides a look at features that relate to the original condition of the meteorite and the significant changes that occurred at later dates. In this technique, a thin slice (about 0.03 millimeter thick) of the meteorite is cut and adhered to a glass plate, thus permitting light to pass through. In this way, researchers can easily identify mineral grains, examine the nature and appearance of any chondrules, and look for evidence of metamorphism and the characteristics of thermal and mechanical alteration resulting from shock impact.

Experimental petrology—laboratory experiments designed to reproduce the mineralogies and textures found in chondrules—has provided valuable data for developing theories to explain chondrule formation and the accretion process of the chondrites themselves. Mathematical models and computer simulations also are used to try to explain chondrite origin. All these different techniques provide clues about the conditions and processes of meteorite formation and their relationship to planetary formation.

CONTEXT

Meteorites in general provided humans' first contact with extraterrestrial materials. Chondrites are similar to certain types of Earth rocks in both their chemistry and their mineralogy. The chondrules they contain have proven to be the oldest known solid material in the solar system. Detailed examination of these chondrules and their matrix material reveals evidence of the processes leading to the formation of the planets.

The carbonaceous chondrites provide information about the organic chemistry that developed in space and the origin of the compounds that eventually may have led to the beginning of life on Earth. The study of chondritic meteorites has revealed much about the conditions that existed when the solar system was formed. They indicate a set of unique conditions that may have lasted only a short period of time and therefore may be the key to understanding how the terrestrial planets came to exist. They have given us new perspectives about our origins and our home, the Earth.

Paul P. Sipiera

FURTHER READING

Bevan, Alex, and John De Laeter. *Meteorites: A Journey Through Space and Time*. Washington, D.C.: Smithsonian Institution Press,

2002. A pictorial work that also deals with radioactive dating and geologic composition of meteorites. Good reference work. For general audiences.

_______. *Thunderstones and Shooting Stars: The Meaning of Meteorites*. Cambridge, Mass.: Harvard University Press, 1986. A very good introduction to the science of meteoritics at a very basic level. It is a good review of the chemical types and methods of study used to classify meteorites. In addition, there is some discussion about the importance of meteorites as a planet-shaping process and about the effect they may have on life-forms throughout the ages. It is best suited for a reading level of high school to college.

Hutchison, Robert. *The Search for Our Beginning: An Enquiry, Based on Meteorite Research, into the Origin of Our Planet and Life*. Oxford, England: Oxford University Press, 1983. A well-written introduction to meteorites and their relationship to planetary formation. Well illustrated; somewhat technical but suitable for high school and college-level readers.

McSween, Harry Y., Jr. *Meteorites and Their Parent Planets*. 2d ed. New York: Cambridge University Press, 1999. Gives an overview of the study of extraterrestrial debris. Good reference for scientists, students, and amateur astronomers.

Mason, Brian. *Meteorites*. New York: John Wiley & Sons, 1962. Perhaps the best of the early books on meteorites and their importance to science. Although this book is dated in the light of modern technology, it still remains an excellent primer for the study of meteorites. It is well written and is best suited for high school and college readers.

Norton, O. Richard. *The Cambridge Encyclopedia of Meteorites*. New York: Cambridge University Press, 2002. Thoroughly discusses interior, external, and atomic compositions of meteorites. Valuable resource for scientists, students, and meteorite enthusiasts.

_______. *Rocks from Space: Meteorites and Meteorite Hunters*. 2d ed. Missoula, Mont.: Mountain Press, 1998. Introduces the reader to meteorites, asteroids, comets, and impact craters. This revised and updated edition includes dozens of new photographs.

Norton, O. Richard, and Lawrence Chitwood. *Field Guide to Meteors and Meteorites*. London: Springer, 2008. This guide is designed to aid those who observe meteors, including locating and analyzing meteorites.

Reynolds, Mike. *Falling Stars: A Guide to Meteors and Meteorites*. Mechanicsburg, Pa.: Stackpole Books, 2001. A guide for nonscientists on observing meteors and meteorites, including how to photograph them and record data.

Spangenburg, Ray, and Kit Moser. *Meteors, Meteorites, and Meteoroids*. Secaucus, N.J.: Franklin Watts, 2002. In addition to covering the basics, this book covers impacts with Earth, how to watch and hunt meteorites, and what can be learned about the universe from them.

Wasson, John T. *Meteorites: Their Record of Early Solar-System History*. New York: W. H. Freeman, 1985. A well-written and well-illustrated introduction to the science of meteoritics. The author covers most of the significant topics in a clear and understandable way and offers a wealth of information for both the casual reader and the serious student. Suitable for high school and college levels.

See also: Asteroids; Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Impact Cratering; Kuiper Belt; Lunar Craters; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Planetary Formation; Pluto and Charon; Solar System: Element Distribution; Solar System: Origins.

Meteorites: Nickel-Irons

Category: Small Bodies

Nickel-iron meteorites (often simply called iron meteorites) are one of the three main groups of meteorites. The nickel-iron group has an approximate composition ratio of more than 80 percent metals to less than 20 percent stony material. The metal in them mostly is iron, with nickel present in much smaller amounts.

OVERVIEW

Meteorites are objects of extraterrestrial origin that have intersected the orbit of the Earth, survived passage through the atmosphere, and reached the Earth's surface in various stages of preservation. Mineralogically, meteorites may contain various proportions of nickel-iron alloys, silicates, sulfides, and various other minor minerals. They are broadly classified into three major groups: nickel-iron meteorites (often called iron meteorites or just irons), stony-iron meteorites (or stony irons), and stony meteorites (or stones). This classification is based on the ratio of metallic to stony minerals. The irons generally contain more than about 80 percent metals, the stony irons have about a 50-50 ratio, and the stones generally contain more than 80 percent stony minerals.

Stony meteorites are by far the most common, accounting for approximately 95 percent of the meteorites observed to fall to Earth and the large number of meteorites that have been collected in Antarctica. Iron meteorites once were thought to be much more common than they actually are, since most meteorites that had been found (not just those observed falling) were irons. This was because iron meteorites are more easily noticed on the ground, standing out more distinctly from terrestrial rocks than the other types do. Collecting meteorites in Antarctica, where all types stand out prominently against a white background of snow and ice, has now shown that irons account for only about 3 to 4 percent of recovered meteorites. Stony irons are even rarer, accounting for no more than 1 percent of the total.

Iron meteorites are composed of nickel-iron alloy minerals that occur in the metallic state. There is no native terrestrial equivalent for these minerals, and in fact the only native metallic iron found on Earth is in small amounts on Disko Island, Greenland, and in Josephine, Oregon. The most common form in which terrestrial iron is found is in the oxide state, in the minerals hematite, magnetite, and limonite. In contrast, the conditions under which meteoritic iron formed were oxygen-poor. This absence of oxygen, combined with the percentage of nickel alloyed with iron, indicates an extraterrestrial origin.

Iron meteorites, or siderites as they were once called, are characterized by the presence of two nickel-iron alloy phases consisting of kamacite ($Fe_{93}Ni_7$) and taenite ($Fe_{65}Ni_{35}$), combined with minor amounts of troilite (FeS) and other rare mineral phases. Based on the percentage of nickel to iron present, iron meteorites are divided into three subgroups: hexahedrites, octahedrites, and ataxites. Hexahedrites possess a bulk chemical composition of 4-6 percent nickel, occurring principally in large single crystals of the mineral kamacite. Octahedrites, which are the most common, contain increasing amounts of nickel, appearing in the mineral form of taenite along with kamacite. The third group, the ataxites, has a nickel content in excess of 18 percent, with taenite and an intergrowth mixture of kamacite and taenite called plessite present.

The two nickel-iron alloy minerals kamacite (up to 7.5 mass percent nickel) and taenite (between 20 and 50 mass percent nickel) are the two most abundant minerals in iron meteorites. More than forty other minerals have also been identified but are present in only minor amounts. Among these minerals, troilite, diamond, and graphite are the most significant. The others have no terrestrial equivalent and have been reported only from meteorite studies.

The mineralogy of iron meteorites is unique also in textural appearance as a result of the relationship between the coexisting kamacite and taenite during the meteorite's cooling process. A mixture of kamacite and taenite produces a geometric pattern of intersecting crystals called Widmanstätten structure, named for its discoverer, Alois Josep Widmanstätten (1754-1849), director of the Imperial Porcelain Works in Vienna. This weave-like or crosshatched pattern is revealed when a cut surface of the iron meteorite is polished and then etched with nitric acid. The pattern results from plates of kamacite occurring in octahedral orientation with the spaces in between filled with taenite. The bandwidth of the pattern depends on the width of the kamacite plates, which varies according to their nickel content. The pattern is thought to be the result of slow cooling over millions of years while the iron resided inside a small asteroid-sized body. The Widmanstätten pattern does not occur in any known terrestrial rock; it is an important criterion in positively identifying a piece of iron as an iron meteorite.

The formation of the three subgroups of iron meteorites is directly related to the amount of nickel originally present, falling temperatures, and the resulting rearrangement of iron and nickel atoms; each subgroup was produced as a certain temperature was passed. The process began as the temperature fell below about 1,700 kelvins (1,400° Celsius), allowing taenite to crystallize. When the temperature dropped below about 1,120 kelvins (850° Celsius), diffusion of nickel occurred, and the crystal structure of the taenite readjusted to accommodate the formation of kamacite. That was possible because both minerals have crystal structures with cubic symmetry, but the size difference between nickel and iron atoms gives each mineral a different crystal form. Kamacite has a "body-centered" crystal lattice; each atom is found at the center of a cube and is surrounded by eight neighboring atoms. In contrast, taenite has a "face-centered" crystal lattice, with an atom centered on each face of a cube; each atom is surrounded by twelve neighboring atoms. The packing arrangement of the atoms in taenite is more compact, thus allowing it to fill the spaces between the kamacite plates.

The study of cooling rates as determined for numerous iron meteorites reveals a wide range. This finding implies that they originated at several different depths rather than in a single core, as once thought. If so, the parent body would have been relatively small (probably between 100 and 300 kilometers in diameter) and would have had a mass insufficient to melt its interior totally. Partial melting could have taken place as a result of radioactive heating as isotopes such as aluminum 26 decayed; this could have created pockets of molten nickel-iron randomly scattered throughout the parent body. Later impacts with similar-sized bodies could have freed them to assume independent orbits as relatively pure lumps of metal alloys. The shock deformation lamellae (called Neumann lines) seen in the hexahedrites may be evidence of such events.

This meteorite, found by the Mars Exploration Rover Opportunity on the Martian surface, was nicknamed Heat Shield Rock and consists mostly of nickel and iron. (NASA)

METHODS OF STUDY

Field recognition of a meteorite is not an easy task, unless one is very familiar with its distinctive characteristics. Usually the most obvious feature of a meteorite will be its unusual heaviness as compared to terrestrial rocks of similar size. This is especially the case for iron meteorites, since they are generally about three times denser than typical Earth rocks. Another easily testable property of an iron meteorite is its strong attraction to a magnet. The surface of a meteorite is fairly smooth and featureless but will often exhibit flowlines, furrows, shallow depressions, and deep cavities. One very characteristic surface feature is shallow depressions known as thumbprints, because they resemble the imprints of thumbs pressed into soft clay. Newly fallen meteorites can also exhibit a fusion crust, which shows the effects of intense atmospheric heating upon its surface. In appearance, this crust resembles black ash, but it will weather to a rusty brown and even disappear with time. The fusion crust on iron meteorites is not particularly distinctive and does weather rapidly.

In most cases, positive confirmation of an iron meteorite must be made in the laboratory. A small corner of the specimen can be cut, polished, and etched with acid to look for Widmanstätten patterns. If these patterns are found, then the specimen is definitely an iron meteorite, but not all iron meteorites show Widmanstätten structure. A relatively simple chemical test for the presence of nickel can be made by dissolving a small amount of the specimen in hydrochloric acid; then tartaric acid, 1 percent solution of dimethylglyoxime in ethanol, and ammonium hydroxide are added. If the solution contains nickel, a scarlet precipitate will result. A quantitative analysis is conducted then to determine the actual mass percentage of nickel. Because nickel content in meteorites falls within a very specific range, this determination will confirm the sample's identity.

Over the years, various criteria have been used to classify iron meteorites. Some of the more obvious have been chemical, structural, and mineralogical; others include cosmic-ray exposure ages and cooling rates. A widely used system, which goes back to the late 1800's, is based on the bandwidth of the Widmanstätten structure (the octahedral array of kamacite) as seen on a cut, polished, and etched surface. The width of these bands of kamacite is dependent on both nickel content and cooling rates. Bulk nickel content generally increases as the bandwidth of kamacite decreases, thus providing a criterion for assigning individual specimens to common groups. Chemical studies for trace elements have extended this classification scheme by including analyses for gallium and germanium. A good correlation has been found between bandwidth size and gallium content, thus permitting a finer separation of iron meteorites into smaller subgroups.

Studies that classify the irons into specific types also provide clues to the meteorite's origin and the nature of its parent body. An estimation of the cooling rate for the coexisting kamacite and taenite can provide evidence of conditions at the time of the meteorite's origin. This cooling rate has been determined from crystallization experiments in the laboratory and from direct observation of the mineral phases found in iron meteorites. The estimated cooling rates vary with the bandwidth sizes of the kamacite phase, and this provides a correlation between cooling rates and bulk chemical composition.

The determination of the cosmic-ray exposure age of a meteorite indicates when the object broke out of its parent body. This technique may also lead to the matching of individual meteorite specimens to a common event. In addition, the compositions and abundances of minor and trace minerals, along with the extent of shock damage to their structures, might give a clearer picture of the events that led to and occurred during the parent body's breakup. Studies such as these reveal clues not only about the origin of the meteorite but also about the formation of the Earth.

CONTEXT

Nickel-iron meteorites had an effect on early human history and technology. Some of the earliest historical records from ancient Egypt speak of iron falling from the sky, and it was undoubtedly meteoritic iron that was first fashioned into iron tools and weapons. Studies have shown that iron tools manufactured on South Pacific Islands, where no local source of iron could be found, were actually forged from meteoritic iron. Some ancient cultures also worshiped "heavenly" iron and placed it in the burial tombs of their leaders; it was thought to be a gift from the gods and served as a symbol of wealth and power. In Europe, as the Bronze Age

About twenty thousand years ago, near Winslow, Arizona, an iron meteorite struck Earth, creating what is now known as Meteor Crater, 200 meters deep and 1 kilometer in diameter. (D. Roddy, U.S. Geological Survey/Lunar and Planetary Institute)

ended, iron actually became more valuable than gold. Perhaps in the not-too-distant future, space colonists will be mining iron asteroids to provide for their industrial needs.

Today, nickel-iron meteorites are helping reveal the processes of planetary formation in the early solar system. They provide evidence of what the interiors of the terrestrial planets may be like. The Earth's core probably is composed of a nickel-iron alloy similar to that found in iron meteorites.

The scars of giant impacts, many due to iron meteorites, dot the Earth's surface from Arizona to Australia. Perhaps the most recent testimony to the effects of a giant meteorite impact can be seen at Meteor Crater near Winslow, Arizona. At that site, more than twenty thousand years ago, an iron meteorite weighing more than 100,000 tons collided with the Earth. The resulting crater, which measures more than 1 kilometer across and nearly 200 meters deep, was created by an object about 30 meters across traveling at a speed of 15 kilometers per second. The energy released at impact was on the order of a 2- or 3-megaton nuclear weapon, destroying most of the meteorite in the process, but that it was an iron meteorite has been confirmed because broken fragments and solidified droplets of meteoritic iron have been recovered around the site.

The largest known intact meteorite is an iron meteorite, the Hoba West meteorite, which weighs an estimated 60 tons and is still embedded in the ground where it fell near Grootfontein in Namibia, southwestern Africa. The largest meteorite on display in a museum is another iron meteorite, the Ahnighito meteorite, which weighs 34 tons; it was found by the arctic explorer R. E. Perry near Cape York, Greenland, in 1894, and brought to New York for display in the American Museum of Natural History.

Paul P. Sipiera

FURTHER READING

Bevan, Alex, and John De Laeter. *Meteorites: A Journey Through Space and Time*. Washington, D.C.: Smithsonian Institution Press, 2002. A pictorial work that also deals with radioactive dating and geologic composition of meteorites. Good reference work. For general audiences.

Buchwald, Vagn F. *Handbook of Iron Meteorites: Their History, Distribution, Composition, and Structure*. Berkeley: University of California Press, 1975. This three-volume work offers both an excellent introduction to the science of meteoritics and a general reference to specific iron meteorites. For the average reader, it provides all the basics on meteorites' origin and chemical nature. For the scientist, it provides the best possible reference source for individual specimens. Most suited for college and graduate levels.

Dodd, Robert T. *Thunderstones and Shooting Stars: The Meaning of Meteorites*. Cambridge, Mass.: Harvard University Press, 1986. A very good introduction to the science of meteoritics at a very basic level. Reviews the chemical types and methods of study used to classify meteorites and makes reference to the importance of meteorites as a planet-shaping process and to the effect they may have had on life-forms throughout the ages. Suitable for high school and college students.

Hutchison, Robert. *The Search for Our Beginning: An Enquiry, Based on Meteorite Research, into the Origin of Our Planet and Life*. Oxford, England: Oxford University Press, 1983. A well-written, well-illustrated introduction to meteorites and their relationship to planetary formation. Although the book is technical, it is understandable for the average reader. Best suited to college-level readers.

McSween, Harry Y., Jr. *Meteorites and Their Parent Planets*. 2d ed. New York: Cambridge University Press, 1999. An overview of the study of extraterrestrial debris. Good reference for scientists, students, and amateur astronomers.

Mason, Brian. *Meteorites*. New York: John Wiley & Sons, 1962. This is perhaps the best of the early books on meteorites and their importance to science. Although the book is dated in the light of modern technology, it remains an excellent primer for the study of meteorites.

Norton, O. Richard. *The Cambridge Encyclopedia of Meteorites*. New York: Cambridge University Press, 2002. Thoroughly discusses interior, external, and atomic compositions of meteorites. Valuable resource for scientists, students, and meteorite enthusiasts.

_______. *Rocks from Space: Meteorites and Meteorite Hunters*. 2d ed. Missoula, Mont.: Mountain Press, 1998. Introduces the reader to meteorites, asteroids, comets, and impact craters. This revised and updated edition includes dozens of new photographs.

Norton, O. Richard, and Lawrence Chitwood. *Field Guide to Meteors and Meteorites*. London: Springer, 2008. A guide to aid readers in observing meteors as well as locating and analyzing meteorites.

Reynolds, Mike. *Falling Stars: A Guide to Meteors and Meteorites*. Mechanicsburg, Pa.: Stackpole Books, 2001. A guide for nonscientists to observing meteors and meteorites, photographing them, and recording data about them.

Sears, D. W. *The Nature and Origin of Meteorites*. Bristol, England: Adam Hilger, 1978. A very readable introduction to the study of meteorites, especially in its historical treatment and its review of the basic concepts. The book does go quite deeply into specialized areas, but it will help the casual reader gain a better perspective on the subject matter. Best suited for the college level.

Spangenburg, Ray, and Kit Moser. *Meteors, Meteorites, and Meteoroids*. Secaucus, N.J.: Franklin Watts, 2002. Information about meteorites, meteors, and meteoroids. This book also discusses impacts with Earth, watching and hunting meteorites, and what can be learned about the universe from them.

Wasson, John T. *Meteorites: Their Record of Early Solar-System History*. New York: W. H. Freeman, 1985. This book is a well-written and well-illustrated introduction to the science of meteoritics. The author covers most of the significant topics in a clear and understandable way and offers a wealth of informa-

tion for both the casual reader and the serious student.

See also: Asteroids; Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Impact Cratering; Kuiper Belt; Lunar Craters; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Planetary Formation; Pluto and Charon; Solar System: Element Distribution; Solar System: Origins.

Meteorites: Stony Irons

Category: Small Bodies

Stony-iron meteorites are intermediate in composition between stony meteorites and iron meteorites. The two major types of stony-iron meteorites are the pallasites and the mesosiderites. The study of pallasites provides evidence for constraints on planetary differentiation processes. The mesosiderites record a history of repeated impacts of projectiles on the basaltic surfaces of their parent body.

OVERVIEW

Meteorites are divided into three broad categories: stony meteorites (or stones), nickel-iron meteorites (or irons), and a group called the stony irons that have both stone and iron components. These stony-iron meteorites are quite rare, constituting only about 1 percent of all the meteorites recovered soon after their fall to Earth was actually observed. The stony irons are more important than their low abundance suggests, however, since they provide a link between the stones and the irons and serve as probes of certain planetary processes. There are four distinct types of stony-iron meteorite: pallasites, mesosiderites, siderophyres, and lodranites. The pallasites and mesosiderites are the most common stony irons; the siderophyres and lodranites are quite rare, represented by only a few specimens each.

Pallasite meteorites are composed of millimeter- to centimeter-sized angular or rounded fragments of magnesium-rich olivine set in a continuous matrix of nickel-iron. In these meteorites, the olivine content ranges from 37 to 85 percent by volume, with the nickel-iron metal accounting for almost all of the remaining material. The minerals troilite, schreibersite, and chromite are sometimes found in small amounts.

The detailed process by which the pallasites formed is still a subject of scientific debate, but they appear to sample a boundary region where nickel-rich iron was in contact with silicate crystals, an environment analogous to the Earth's core-mantle boundary. One mechanism for the formation of pallasites could have been the heating and consequent differentiation of a chondritic parent body. The high-density iron-nickel-sulfur liquid settled to the center, forming a molten core and leaving a silicate-rich mantle. As the mantle cooled, olivine, which is generally the first silicate mineral to crystallize out of cooling silicate liquids of a wide range of compositions, formed and settled to the core-mantle boundary.

The mechanism by which molten metal from the core surrounded the olivine crystals to produce the pallasite structure is not yet understood. It has been proposed that perhaps the mantle shrank as it cooled, squeezing molten metal out of the core and into the olivine-rich layer. Alternatively, the core may have contracted during cooling, causing the olivine layer to collapse into the void, giving rise to the mixing. Further cooling would have resulted in the solid pallasite material, which later was excavated from the parent body by major impacts. Therefore, the pallasites are thought to provide samples similar to the core-mantle boundary region on Earth.

Comparison of the chemistry of the pallasites with that of the Earth provides some constraints on the Earth's formation and differentiation process. The Earth is generally assumed to have formed with the same chondritic composition as the pallasite parent body. After differentiation, the concentration of nickel in the Earth's upper mantle remained at about 0.2 percent. The silicates in the pallasites are much more depleted in nickel, having a concentration of only 0.002 percent. One possible explanation for the additional nickel in the Earth's outer layers is that after differentiation, additional chondritic material was added to the surface, presumably by impacting objects.

Constraints on the size of the pallasite parent body come from a study of how fast these objects cooled after differentiation. If two objects start at the same temperature and are allowed to cool, the smaller object will cool more rapidly, since it has a larger ratio of surface area to volume than does the larger object. The cooling rates determined for the pallasites, and the iron meteorites related to them, are consistent with formation in an object much smaller than the Earth's moon, perhaps no larger than 10 kilometers in diameter. The texture, composition, and cooling rate of typical pallasites are consistent with their metal being related to a group of iron meteorites called the IIIAB irons. If so, then samples of the pure core material of the pallasite parent body are also available as the IIIAB irons.

The differentiation process believed to have occurred in the early history of the Earth and of the pallasite parent body has been simulated in the laboratory by heating chondritic meteorites. As the temperature increases, the meteorites melt in stages. The first liquid to appear is composed mainly of iron, nickel, sulfur, and trace elements that have an affinity for these major elements. Because this liquid is twice as dense as the remaining silicates, it sinks to the bottom. Further melting yields liquids of basaltic composition and a solid residue of mostly olivine. The basaltic liquid, which is less dense than the solids, floats to the top. When cooled, the resulting structure has metal at the bottom, an olivine layer in the middle, and basaltic material on top. For the Earth, this process would give rise to a dense metal core surrounded by an olivine-rich mantle and covered with a basaltic crust. The absence of samples from the Earth's deep interior, however, prevents direct verification of this structure.

Examination of the pallasite meteorites strongly suggests that the pallasite parent body formed with a chondritic composition, was heated and melted, differentiated into a metallic core and silicate mantle, and subsequently cooled and solidified. Thus, the pallasites confirm that planetary differentiation took place on the pallasite parent body in the same manner as proposed for the Earth.

The mesosiderite meteorites are quite different from the pallasites. They are composed of angular chunks of basaltic rocks and rounded masses of metal. The metal phases constitute 17 to 80 percent of the mesosiderites by weight. The major silicate minerals are plagioclase feldspar, calcium-rich pyroxene, and olivine. The mesosiderites are polymict breccias; that is, they are composed of fragments of unrelated rocks. They contain pyroxene-rich fragments, like the diogenite achondrite meteorites, and fine-grained fragments of eucrite achondrite meteorites. The eucrites and diogenites are composed of magmatic rocks similar, respectively, to terrestrial basalts and cumulates.

The mesosiderites appear to have formed from repeated impacts on an asteroidal surface, which brought together at least three distinct types of material: diogenitic and eucritic rocks from the surface of the asteroid and a nonindigenous metallic component, possibly from the impacting objects. If the metal fragments in the mesosiderites are projectile material from the core of a previously fragmented asteroid, these fragments must have struck the surface of the diogenite-eucrite parent body at a very low velocity. Impacts at velocities higher than about 1 kilometer per second lead to very low concentrations of the projectile material in the resulting breccias. This low-impact velocity would suggest that the parent body exerted a very small gravitational attraction on the falling metal, indicating that the diogenite-eucrite parent was a relatively small asteroid, not a planet-sized object.

The mesosiderites are similar to lunar surface breccias, which also formed by multiple impacts into basaltic rock. They allow the processes of basaltic volcanism and impact brecciation to be examined in a different solar system region and an earlier time than occurred on the Moon.

There are a few other meteorites that contain mixtures of metal and silicate phases, but they are otherwise dissimilar to the pallasites and mesosiderites. The siderophyre type, represented only by the single meteorite known as Steinback, is composed of the silicate mineral bronzite (an iron-magnesium pyroxene) and metal. The lodranites are composed of olivine, calcium-poor pyroxene, and metal. These two rare types of stony-iron meteorite have not been as well studied as the pallasites and mesosiderites.

METHODS OF STUDY

The pallasite meteorites have been well studied by a variety of techniques, because, along with the iron meteorites, they provide a window on the processes and conditions in the deep interior of their parent bodies and clues to similar processes thought to have occurred on Earth. Much of the evidence concerning these processes comes from detailed analyses of the chemical abundances of major and trace elements in individual minerals from each meteorite. Detailed modeling of the differentiation process suggests that certain metal-seeking trace elements will concentrate in the metallic core, while other trace elements concentrate in the silicate mantle.

Early studies of the metal phases in iron and stony-iron meteorites were done by examining their textures, because the abundances of the trace elements were difficult to determine. More recently, however, the abundances of these trace elements, present at the level of no more than a few atoms in every million atoms of bulk material, have been measured by neutron activation, X-ray fluorescence, and electron microprobe analysis.

Detailed measurements of the abundance of trace elements, emphasizing the elements gallium, germanium, and iridium, in iron meteorites, show twelve to sixteen distinct compositional clusters, indicating that the irons sample a minimum of twelve different parent bodies. Almost all the pallasites have metal compositions and textures, suggesting that they are related to a single group of iron meteorites, the IIIAB irons. This relationship suggests that the metal portion of the pallasites samples the core of the same parent body as the IIIAB irons.

Chemical analysis of the olivine grains in pallasites indicates that the olivine has a very narrow range of compositions. Within each meteorite, the olivine crystals are homogeneous; that is, they show no significant compositional variation from grain to grain. This narrow range of olivine compositions suggests that the grains formed from a silicate liquid of uniform composition. Most of the pallasites, then, appear to sample the core-mantle boundary of a single parent body.

A few pallasites differ from the majority in that they contain olivine richer in iron than the common pallasites. A few of these pallasites are also enriched in nickel and the trace elements germanium and iridium and depleted in gallium relative to the common pallasites. The trace element abundances suggest that these pallasites sample a parent body different from the common pallasites; however, the metal in these cannot be identified with any iron meteorite group.

The cooling rates of pallasite meteorites, from which the size of the parent bodies can be inferred, are determined by examination of the metal. Meteoritic metal consists of two distinct nickel-iron alloys: kamacite, an alloy that can be no more than 7 percent nickel, and taenite, which frequently has more than 20 percent nickel. Metallic liquid cores generally are thought to have a higher nickel content than can be accommodated in the kamacite structure alone. As the metal cools, the amount of kamacite increases, and nickel atoms diffuse from the newly formed kamacite into the nearby taenite. Since nickel diffuses more rapidly in kamacite than in taenite, however, the nickel will build up at the kamacite-taenite boundaries. There will therefore be more nickel at the edges of the taenite than near the center. This distribution of nickel in the taenite varies with the cooling rate.

Electron microprobe analysis of the taenite grains gives the abundance of nickel as a function of distance from the edge, which allows the cooling rate to be estimated. The nickel distribution in the metal of the normal pallasites is consistent with a very rapid cooling rate, implying an extremely small parent body (less than 10 kilometers in diameter). The same technique suggests that the cooling rate of the IIIAB iron meteorites, with which the pallasites are apparently associated, was somewhat slower, implying a parent body of 200 to 300 kilometers in diameter. The reason for this difference is not yet understood.

The mesosiderites, although they are also mixtures of stone and metal, are quite different from the pallasites. The silicate portion of the mesosiderites is rich in the minerals plagioclase feldspar and calcium-rich pyroxene. These minerals melt at relatively low temperatures and are common on the surface of the Earth and the Moon. Unlike the olivine found in the pallasites, the basaltic minerals found in the mesosiderites were probably never in direct contact with the metal in the core of the parent body.

Detailed examination of the mineralogy of the mesosiderites shows that they consist of a mixture of three distinct components, each represented by a distinct type of meteorite. The silicates are fragments of both eucrite and diogenite achondrite stony meteorites; the metal phases resemble the iron meteorites. The eucrites are basaltlike meteorites composed mainly of calcium-poor pyroxene and plagioclase feldspar thought to have crystallized on or near the surface of their parent body. The diogenites are composed mainly of bronzite (an iron-magnesium pyroxene) which resembles the pyroxene cumulates found in the Stillwater complex and in other layered terrestrial intrusions. This combination of eucrite and diogenite fragments in the stony-iron mesosiderites, and in the howardite achondrite stony meteorites, is taken as evidence that the eucrites and diogenites formed on the same parent body. The metal in the mesosiderites occurs mainly in the form of nuggets, sometimes up to 9 centimeters in diameter, or fragments. They are quite distinct from the metal veins that are continuous throughout the pallasites. Trace element analysis of the metal in the mesosiderites suggests a similarity with the IIE iron meteorites; however, the link is much weaker than that of the common pallasites with the IIIAB iron meteorites.

CONTEXT

The process of differentiation (the melting of a primitive parent body and the concentration of iron-nickel-sulfur in the core and the lighter silicate minerals in the mantle) is believed to have taken place on Earth. The study of the pallasite meteorites, which are composed of a mixture of iron-nickel core material and olivine mantle material, confirms that this process of differentiation did occur early in solar-system history, at least on the pallasite parent body. Though the chemical compositions of most pallasites are consistent with a single parent body, the few pallasites of unusual composition suggest that the pallasite meteorites sample core-mantle boundaries of several parent bodies. That indicates that the differentiation process was relatively common. The good match between the postulated chemical compositions of the core and mantle and the compositions actually seen in the metal and olivine phases of the pallasites confirms the model of chemical segregation developed for the differentiation of the Earth. It is believed, however, that the common pallasites sample the core-mantle boundary of a much smaller body than the Earth.

The mesosiderites, although they also consist of a mixture of metal and silicates, are quite different from the pallasites. The silicates in the mesosiderites apparently sample basaltic material similar to that found in the lunar surface, in the Earth's crust, and, as implied by chemical measurements taken by spacecraft, on the surfaces of Mars and Venus. The iron fragments may be projectiles that struck the rocky surface of the mesosiderite parent body and were incorporated into the rock produced by the impact. The mesosiderites demonstrate that basaltic rocks similar to those on Earth occur on the surfaces of some asteroids and that the impact processes that dominate the lunar landscape also occurred in the early history of the solar system.

George J. Flynn

FURTHER READING

Bevan, Alex, and John De Laeter. *Meteorites: A Journey Through Space and Time*. Washington, D.C.: Smithsonian Institution Press, 2002. A pictorial work that also deals with radioactive dating and geologic composition of meteorites. Good for general audiences.

Dodd, Robert T. *Thunderstones and Shooting Stars: The Meaning of Meteorites*. Cambridge, Mass.: Harvard University Press, 1986. This reference book explains why there is a scientific interest in meteorites and summarizes what is known about them. It includes chapters on why planets melt and on the iron and pallasite meteorites. Although written as a college-level text, it provides detailed explanations of the phenomena and techniques of analysis that will be understood by the reader without an Earth science background.

Kerridge, John F., and Mildred S. Matthews, eds. *Meteorites and the Early Solar System*. Tucson: University of Arizona Press, 1988. A collection of articles by sixty-nine contributing authors describing the state of meteorite research as it relates to early solar system processes. Section 3.2 focuses on igneous activity and the process of differentiation on the parent bodies of the iron and stony-iron meteorites. Aimed at graduate students studying meteoritics, with a detailed review and comprehensive bibliography of the major topics in meteorite research. Well illustrated.

McSween, Harry Y., Jr. *Meteorites and Their Parent Planets*. 2d ed. New York: Cambridge University Press, 1999. An overview of the study of extraterrestrial debris. Good reference for scientists, students, and amateur astronomers.

Norton, O. Richard. *The Cambridge Encyclopedia of Meteorites*. New York: Cambridge University Press, 2002. Thoroughly discusses interior, external, and atomic compositions of meteorites. Valuable resource for scientists, students, and meteorite enthusiasts.

_______. *Rocks from Space: Meteorites and Meteorite Hunters*. 2d ed. Missoula, Mont.: Mountain Press, 1998. Introduces the reader to meteorites, asteroids, comets, and impact craters. This revised and updated edition includes dozens of new photographs.

Norton, O. Richard, and Lawrence Chitwood. *Field Guide to Meteors and Meteorites*. London: Springer, 2008. This guide is designed to aid readers in observing meteors, as well as locating and analyzing meteorites.

Reynolds, Mike. *Falling Stars: A Guide to Meteors and Meteorites*. Mechanicsburg, Pa.: Stackpole Books, 2001. A guide to observing meteors and meteorites, photographing them, and recording data. For nonscientists.

Spangenburg, Ray, and Kit Moser. *Meteors, Meteorites, and Meteoroids*. Secaucus, N.J.: Franklin Watts, 2002. Covers impacts with Earth, watching and hunting meteorites, and what can be learned about the universe from them.

See also: Asteroids; Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Impact Cratering; Kuiper Belt; Lunar Craters; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Oort Cloud; Planetary Formation; Pluto and Charon; Solar System: Element Distribution; Solar System: Origins.

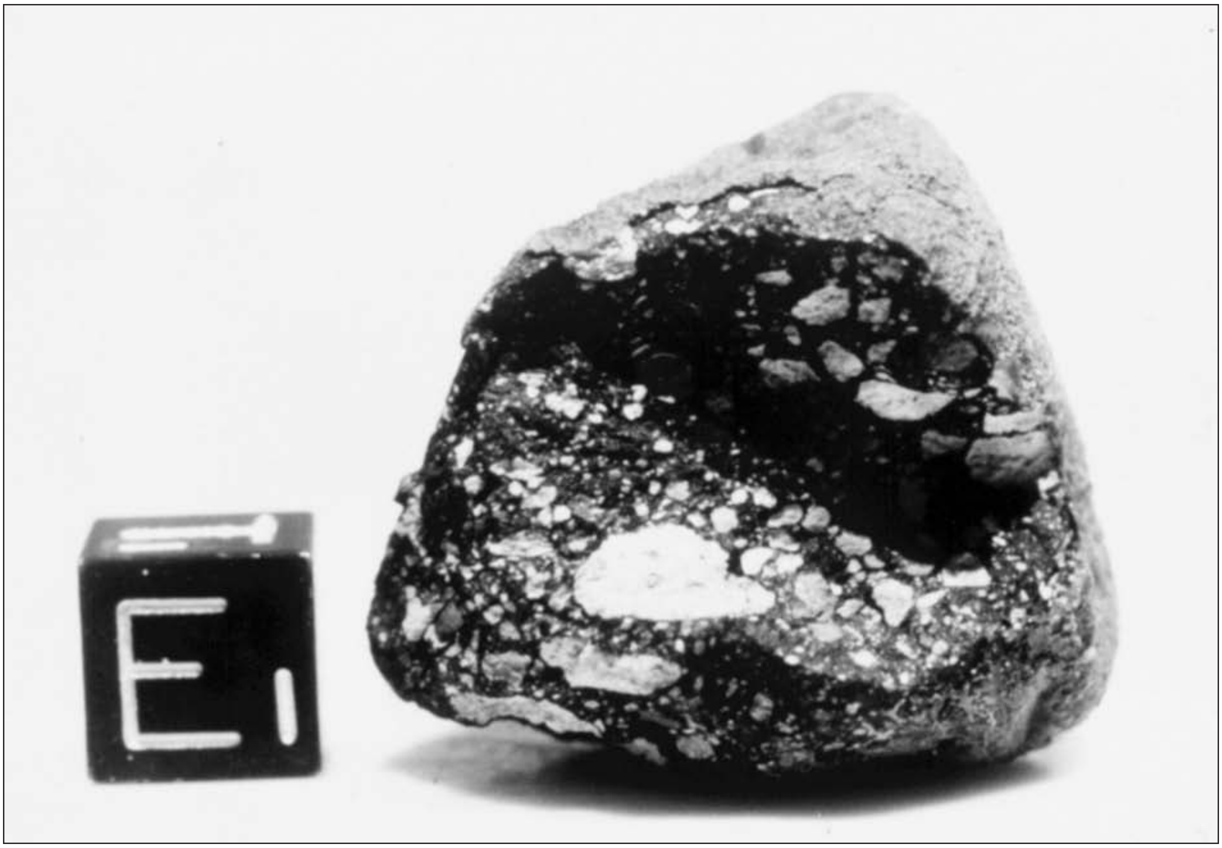
Meteoroids from the Moon and Mars

Categories: Mars; Small Bodies

Prior to the 1969 Apollo lunar landings, meteorites represented the only extraterrestrial material available for scientific study. In 1979 scientists found an unusual meteorite in Antarctica that resembled a certain type of moon rock. Numerous comparative studies confirmed that this Antarctic meteorite was of lunar origin and the result of an impact event. Subsequent studies on a small group of anomalous meteorites provided evidence to support their Martian origin.

OVERVIEW

Throughout human history people have been aware of rocks falling from the sky. In earlier times they were thought to be either gifts from the gods or the result of severe weather picking up rocks from one location and then dropping them at another. Later, as human curiosity developed into science, people began to look differ-



Found in Antarctica in 1981, this meteorite—which includes white fragments of anorthite, typical of the lunar highlands—was among the first to show that not all meteorites originate in the asteroid belt. (NASA/Johnson Space Center)

ently at these strange rocks and associated the appearance of a bright fireball with the fall of a meteorite. They also noted that meteorites were composed of different substances. Some were made of nearly pure metal, while others had the appearance of some volcanic rocks. Scientists later performed chemical and mineralogical analyses that identified elements and minerals in meteorites that were either rare or unknown on Earth. Eventually all this evidence led to the conclusion that meteorites originated in space and were not from the Earth.

Meteorites were once the only extraterrestrial material available for scientific study. That changed when the six Apollo lunar landings returned to Earth more than 383 kilograms of rocks and other lunar samples. Later, Soviet Luna missions returned a small sampling of lunar material. Among the rock specimens collected were examples of breccia, basalt, anorthosite, and gabbro. Along with regolith samples, this sampling of lunar rocks has provided scientists with enough information to gain a good understanding of the Moon's composition, internal structure, and geological history throughout time. However, Apollo samples were limited to only six locations, and a couple of these sites were very similar to each other in their geological settings. Vast areas of the lunar surface, including the lunar far side, were not represented. Scientists were anxious to get their hands on rock specimens from these unexplored regions. Little did they know that meteorites would once again provide the required research material.

Along with the 1969 Apollo lunar landings, another rather important event took place in planetary science that year. In Antarctica, a team of Japanese scientists working in the Yamato Mountains region came across nine meteorites sitting on the ice in close proximity to one another. Later, after these meteorites had been classified, it was determined that they represented several different falls. This was extraordinary and suggested that Antarctica might be a "treasure house" for meteorite finds. Subsequent search teams have recovered tens of thousands of meteorites, confirming the theory that Antarctica offers a unique situation for meteorite finds. Found among this huge number of specimens is the full range of stone, stony-iron, and iron meteorite types. Also present are a relatively large number of the biologically important carbonaceous chondrites and a variety of achondrites.

During the 1979-1980 Antarctic field season, another group of Japanese scientists searching for meteorites in the Yamato Mountains recovered what was at first believed to be an anomalous achondrite. Upon later examination, researchers recognized the appearance of a lunar breccia. After additional analyses and comparison to Apollo lunar rocks, this specimen was determined to be a lunar meteorite. It had been blasted off the Moon's surface during a large impact event and was later pulled to Earth and entered the atmosphere just like any other meteorite. Confirming that this specimen came from the Moon could not have been possible without the availability of the Apollo lunar specimens for comparison.

Confirmation that lunar rocks have reached the surface of Earth as meteorites excited the scientific community. Not only did these finds provide additional lunar material for study; they also represented material from areas on the Moon that had not been explored by the Apollo astronauts. Each year, as more lunar meteorites were being found in Antarctica, meteorite dealers and collectors from around the world began to search other locations to find lunar meteorites. Soon the deserts of Australia, Northwest Africa, Libya, and Oman became prime search areas for meteorites. Within a relatively short period, meteorites, found by the local inhabitants, began turning up in marketplaces and were quickly purchased by mineral dealers. Gradually these meteorites found their way to the scientists for study. Found among the huge numbers of ordinary meteorites were the occasional gems, the lunar meteorites and the equally exciting SNC meteorites.

A meteorite from Mars, found in North Africa, is one of only a few Martian meteorites known. (Bruno Fectau and Carine Bidaut)

Prior to the discovery of lunar meteorites in Antarctica, a small group of achondrite meteorites called SNCs (pronounced "snick" and standing for shergottites, nakhlites, and chassignites), could not be readily explained. They were different from most meteorites because of their relatively young radiometric ages and the fact that they were essentially volcanic rocks. When compared to the isotopic chemistry of Earth or Moon rocks, they were clearly different and their origin was unknown. At this time it was generally believed that all meteorites came from the asteroid belt as a result of numerous impact events early in the history of the solar system. This clearly was not the case for the SNCs.

As early as 1979, based on the new understanding of lunar meteorites, some scientists suggested that the SNCs might have come from Mars. Initially most scientists did not readily accept this theory, but that quickly changed. A couple of the suspected Martian meteorites contained glass in which a tiny amount of gas had been trapped long ago, while the rock was on Mars. When scientists analyzed this gas it was found to have a nitrogen and noble gas content that closely matched the gases found in the Martian atmosphere, as sampled by the Viking Landers in 1976. Also, many geologists believed that the massive volcanoes on Mars such as Olympus Mons could have been active at the time that these meteorites were believed to be on Mars. When scientists connected the relatively young age of the SNC meteorites to the volcanoes, it seemed to make sense that the SNCs might be of Martian origin.

To many scientists it seemed quite clear that the best place of origin for these intriguing SNC meteorites had to be Mars. For others, the evidence was not as convincing. The main argument against a Martian origin for the SNC meteorites rests on the fact that scientists do not have a documented sample of a Martian rock for comparison. No astronauts or robotic explorers have collected samples from the Martian surface and returned them to Earth. With no definitive Martian rocks for comparison, the SNCs cannot be verified as truly Martian. The robot explorers Pathfinder, Spirit, and Opportunity have traveled across Mars and viewed hundreds of rocks close up, but they have not seen a type of rock that matches any of the SNC meteorites. Although this lack of evidence does not preclude the possibility that SNCs are of Martian origin, it will take a major sampling effort by both robotic and human explorers to make that determination, and therefore the debate over the origin of the SNC meteorites will continue to spark scientific discussions for years if not decades to come.

KNOWLEDGE GAINED

The basic study of meteorites paved the way for the development of the scientific technology that was required for the examination of materials brought back from the Moon by the Apollo astronauts. Geologists and geochemists studying meteorites have a more difficult task than those who study Earth rocks. The average meteorite fall usually consists of a relatively small amount of material. Proper classification and additional detailed studies require destructive analyses. With only a small amount of material available, scientists had to develop analytical techniques that require only minimal amounts of the precious material available to them. Instruments such as the electron microscope, electron and ion microprobes, and mass spectrometer gave the researcher tools to explore the secrets of the universe from the chemical composition of a meteorite. As is usually the case, technology specifically developed for pure scientific research was immediately integrated into the commercial market and became valuable tools for industry and medical applications.

The rock and regolith samples returned by the six Apollo landings were not representative of the entire Moon and its major geological terrains. Many questions concerning the Moon's origin and relationship to Earth could not be answered with only the Apollo material. Scientists would have to wait patiently for some future missions to bring back samples from different areas. The 1979 find of lunar meteorites in Antarctica changed all that, giving scientists the opportunity to study additional lunar materials. A lunar meteorite represents material blasted off the Moon's surface during an impact event; hence, it does not represent a single location but could have originated anywhere on the lunar surface. By examining lunar meteorites that originated from different locations on the Moon, scientists have been able to construct a more comprehensive picture of lunar geology.

The case for Martian meteorites proved to be equally exciting, especially with the meteorite ALH 84001. This unique meteorite was found during the 1984-1985 field season in the Allan Hills region of Antarctica. Initially classified as an anomalous achondrite, it was later reclassified as a SNC Martian meteorite. Then in 1996, after years of intensive study, scientists at the National Aeronautics and Space Administration (NASA) announced the discovery of possible microscopic fossil bacteria within carbonate deposits present in the meteorite. This revelation created much excitement within the scientific community and sparked intense controversy and debate. Additional studies provided evidence that the features in the meteorite resembling bacteria could have been the result of an inorganic process, too. Neither side in the debate could definitely prove the other wrong, so in the end the "fossil" was deemed inclusive with regard to the evidence of organic material, and scientists concluded that additional evidence was required.

Although the argument for fossil bacteria in ALH 84001 was not fully accepted, it inspired other scientists to look for microscopic evidence of life in more promising meteorites, notably the carbonaceous chondrites. The recognition of lunar and Martian meteorites also suggests that planets can exchange material with each other and possibly distribute the essential chemical compounds of life through giant impacts.

CONTEXT

Confirmation of lunar meteorites led to the possibility of meteorites coming from Mars. Once scientists became comfortable with the idea that rocks could be blasted off the Moon's surface and pulled to Earth, they expanded their options to include other worlds. Because of its relative closeness to Earth and similar geological features, Mars became the logical first choice. Using the same technology developed for lunar studies, scientists began to look at a number of intriguing anomalous meteorites (SNCs) with uncertain places of origin. By comparing gas compositions and isotopic data from these meteorites to data obtained through the Viking Mars Project, it was concluded that they could have come from Mars in a manner similar to the way lunar meteorites arrived. Although this Mars-origin theory for these meteorites is not conclusive, most scientists seem to have accepted it as fact and moved on from there. Now other anomalous types of meteorite are being examined for evidence that may have them coming from Mercury or possibly Venus. Perhaps one day a particularly interesting meteorite may produce an ancient piece of Earth rock tossed out by an early giant impact event that has finally returned home.

Paul P. Sipiera

FURTHER READING

Bevan, Alex, and John De Laeter. *Meteorites: A Journey Through Space and Time*. Washington, D.C.: Smithsonian Institution Press, 2002. A beautifully illustrated book that covers the topic of meteorites at a level that is suitable for a wide readership.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. This book deals principally with the geological aspects of the entire solar system as related to planetary formation. An excellent resource for students of astronomy and geology.

Kring, David A. "Unlocking the Solar System's Past." *Astronomy* 34, no. 8 (August, 2006): 32-37. A good introduction to the subject of meteorites and how they relate to the other members of the solar system. Suitable for general readers.

Lauretta, Dante S., and Harry Y. McSween, eds. *Meteorites and the Early Solar System II*. Tucson: University of Arizona Press, 2006. A comprehensive collection of scientific papers covering the role meteorites play in planetary formation. This book serves a valuable resource for the graduate student and the professional researcher alike.

Norton, O. Richard. *The Cambridge Encyclopedia of Meteorites*. Cambridge, England: Cambridge University Press, 2002. A comprehensive and well-illustrated work on the topic of meteorites. Suited for a wide audience, from high school through graduate students and meteorite enthusiasts.

See also: Asteroids; Mars: Illusions and Conspiracies; Mars: Possible Life; Mars: Surface Experiments; Mars's Atmosphere; Mars's Craters; Mars's Polar Caps; Mars's Satellites; Mars's Valleys; Mars's Volcanoes; Mars's Water.

Meteors and Meteor Showers

Category: Small Bodies

Meteors are those streaks of light produced by small solar-system bodies (meteoroids) entering the Earth's atmosphere. Fragments from asteroids produce sporadic meteors, while debris left along the orbit of a comet causes meteor showers. Both provide information about the origins of the solar system, especially if they reach the ground and are recovered as meteorites.

OVERVIEW

Scientific study of meteors and their relation to meteorites did not start until the beginning of the nineteenth century. Earlier meteorite falls were observed, with stones recovered, but most witnesses were ridiculed, and "sky stones" were treated with suspicion. In the Bible, Joshua 10:11 records a battle in which the enemy was defeated by "stones from heaven," which may have been meteorites. Acts 20:35 refers to the image of Diana of Ephesus standing on a stone that fell from heaven.

Anaxagoras, Plutarch, and several Chinese recorders from as early as 644 B.C.E. described stones falling from the sky. A stone preserved in a corner of the Kaaba in Mecca fell in the seventh century. The oldest authenticated meteorite in Europe, a 120-kilogram stone that fell in Switzerland in 1492, is still preserved in a museum. In spite of this evidence, much doubt remained among scientists in Europe. When a stone fell near Luce in France in 1768, it was studied by French chemist Antoine Lavoisier and two other French scientists, who concluded that it was an ordinary stone struck by lightning.

In 1794, Czech acoustic scientist Ernst Chladni published an account of numerous reported meteorite falls, giving strong evidence that some of them must be of extraterrestrial origin. He stated that the flight of such an object through the atmosphere caused the bright, luminous phenomenon known as a fireball. Chladni found few supporters for this idea, as most held to Aristotle's view that comets and flashes of light across the sky were atmospheric phenomena (the word "meteor" comes from the Greek word for things related to the atmosphere, as in meteorology). Chladni's cosmic theory of meteors was finally confirmed in 1798 by two students at the University of Göttingen, H. W. Brandes and J. F. Benzenberg, who had read his book. They made simultaneous observations of "shooting stars" from two different locations separated by several kilometers and used a simple triangulation method to show that the light flashes originated at least 80 kilometers above the ground from objects moving several kilometers per second from a source beyond the atmosphere. Most doubts about meteorite falls were removed after the physicist Jean-Baptiste Biot reported an unusual fall of two or three thousand stones at L'Aigle in 1803, which eyewitnesses said was preceded by a rapidly moving fireball and explosion.

On a clear, dark night, a diligent observer may be able to witness on average six meteors per hour. More are visible after midnight than before, increasing to a maximum just before dawn. In the 1860's, Italian astronomer Giovanni Schiaparelli, famous for his discovery of Martian *canali* (channels), explained the increase in meteors at certain times as resulting from the Earth's orbital and rotational motion. Before midnight, the observer is on the trailing side of the Earth's motion and can see only those meteors that overtake the Earth. After midnight, an observer is on the leading side of the Earth's motion and will intercept meteors in front of it. Thus, meteors will appear brighter because they are entering the Earth's atmosphere at a higher velocity. Because the Earth's orbital velocity is about 30 kilometers per second, and the escape velocity from the Sun at the Earth's orbital distance is about 43 kilometers per second, solar-system objects should range in speed from 13 to 73 kilometers per second. Because no meteors have been observed with a faster speed, it is believed that they come from within the solar system rather than from interstellar space.

Most meteors become visible about 100 kilometers above the Earth's surface and are completely consumed when they reach about 70 ki-

The Leonid meteor shower, in an image captured by an aircraft on November 17, 1999. The Leonids occur when Earth passes through the debris field of Comet Tempel-Tuttle. (NASA/SAS/Shinsuke Abe and Hajime Yano)

lometers, although a few larger ones reach about 50 kilometers. Most meteors range in size from a few microns up to several millimeters. Survey estimates indicate that about 25 million meteors are bright enough to be seen over the entire Earth in any twenty-four-hour period. Telescopic surveys suggest that several billion meteoroids enter the Earth's atmosphere every twenty-four hours, with an average total mass of about 100,000 kilograms. Most of this is consumed in the atmosphere as meteoroids are heated by friction to incandescence, but on average about 1,000 kilograms per day are deposited on the Earth as meteorites.

More than half of all meteors are called sporadic because they appear at any time and from any direction in the sky. The remaining meteors are associated with meteor showers that appear to radiate from a common point in the sky, called the radiant; they actually move along parallel paths but appear to diverge from the radiant—much like the divergence of railroad tracks when viewed in perspective. Meteor showers recur on an annual basis, with about a tenfold increase over the usual sporadic rate. They are named for the constellation in which the radiant appears to be located. Annual showers occur when the Earth crosses a meteroid stream that fills the orbit of a comet, while periodic showers occur less frequently, when Earth crosses a meteroid swarm in the wake of a comet. The most spectacular periodic meteor showers are the Leonids, whose radiant is located in the constellation Leo. Historical records as far back as 902 C.E. mention the Leonids. A spectacular display on October 14, 934, is described in Chinese, European, and Arabic chronicles. The Japanese recorded a six-hour display in 967, and Chinese records continued to describe them every thirty-three years for several centuries.

The modern study of meteor showers began with the famous naturalist Baron Alexander von Humboldt. He observed the Leonids by chance during a trip to South America in 1799 in a two-hour display of hundreds of thousands of meteors. Humboldt was the first to suggest that these meteors might originate from a common point in the sky. The greatest Leonid display in the nineteenth century was observed in the United States and Canada on November 12, 1833. About one thousand meteors per minute were counted, and the appearance of the radiant was confirmed. The following year, two Americans, D. Olmstead and A. C. Twining, suggested that the annual Leonids were caused by the Earth passing through a cloud of meteoroids each November. A few years later, German astronomer Heinrich Wilhelm Olbers proposed that the more intense periodic meteor showers of 1799 and 1833 were caused by a denser swarm of the Leonid meteoroid stream. In 1864, H. A. Newton of Yale College reached the same conclusion independently and showed a period of recurrence of just over thirty-three years from historical records, beginning with the shower of 902. Their prediction of a spectacular display in 1866 was confirmed. Later, English astronomer John Couch Adams, who theorized the existence of the planet Neptune, succeeded in computing the Leonid stream orbit.

In the 1860's, other meteoroid streams were identified and traced back through history. Records back to the tenth century in England recorded meteor showers associated with the festival of St. Lawrence (August 10), known as "the tears of St. Lawrence" but now identified as the August Perseids from their radiant in Perseus. In 1861, the American astronomer Daniel Kirkwood, who later discovered gaps in the asteroid belt, suggested that meteor showers result from debris left in the wake of a comet through which the Earth occasionally passes. In 1866, Schiaparelli announced that the August Perseids appear to occupy the same orbit as Comet Swift-Tuttle (1862 III). Soon after, Urbain Le Verrier and C. A. F. Peters identified the November Leonids with Comet Tempel-Tuttle (1866 I), which had a recurrence period of thirty-three years. Both the May Aquarids and the October Orionids have been associated with Halley's comet. The greatest naked-eye meteor observer was W. F. Denning, who published a catalog in 1899 of several thousand radiants, mostly of minor meteor showers of less than 10 meteors per hour, based on more than twenty years of observation.

Like comets, meteor streams may be perturbed by planets into new orbits. Those with high inclinations to the ecliptic plane (the plane of the Earth's orbit) or in retrograde orbits (opposite to the Earth's motion) are least affected, such as the Leonids, Perseids, and Lyrids. After the Leonid display of 1866, the main body of the stream passed close to Jupiter and Saturn. Its associated comet could no longer be found, and only a few meteors were observed in 1899 and 1933. The comet was found again in 1965, and then, on the morning of November 17, 1966, the Leonids returned, with meteors as bright as Venus. Viewed from the western United States, they reached a maximum rate of more than two thousand per minute before dawn, producing the greatest meteor display in recorded history.

Only the bright fireball meteors, sometimes brighter than the full Moon, are produced by meteoroids large enough to survive passage through the Earth's atmosphere and fall to the ground as meteorites. Almost all of these are sporadic meteors; even among the fireballs, less than 1 percent yield meteorites. Dozens of meteorites fall to the surface of the Earth each day, but very few are recovered. About 95 percent of "falls" (seen falling and then recovered) are classified as "stones" (about 75 percent silicates and 25 percent iron), but 65 percent of "finds" (whose associated meteors are not observed) are "irons" (90 percent iron and 8 percent nickel), because irons are easier than stones to identify on the ground as meteorites.

Dozens of craters apparently formed by large meteorites have been identified around the world. The first such identification was made about 1900 by Daniel Barringer at Canyon Diablo in Arizona. This crater is 1.3 kilometers across and 180 meters deep, with a rim rising 45 meters above the surrounding plain. About 25,000 kilograms of iron meteorite fragments have been found in and around the crater. It is estimated that the crater was formed by an explosive impact about 50,000 years ago from a 60-million-kilogram meteorite. In 1908, a brilliant fireball meteor exploded in the Tunguska region of Siberia, leveling trees over a distance of 30 kilometers and killing some 1,500 reindeer. No large crater or meteorite has been found, but its effects were estimated to be equivalent to the explosion of a billion-kilogram meteroid. In 1972, a fireball meteor with an estimated mass of a million kilograms was photographed in daylight some 60 kilometers above the Grand Teton Mountains before leaving the atmosphere over Canada.

METHODS OF STUDY

Information about meteors can be obtained with the unaided eye, but much greater scope and precision results from the use of photographic, radar (radio echo), and space-probe techniques. Modern photographic meteor observations were begun by Fred Whipple in 1936 at the Harvard College Observatory using short-focal-length, wide-angle cameras. These were later replaced by ultra-fast Super-Schmidt cameras that could detect meteors as small as a milligram. To measure the height, direction, and velocity of a meteor, simultaneous photographs of the meteor trail are taken from two stations separated by about 50 kilometers. Each photograph shows the positions of the meteor trail against the background of stars from each station so that its trajectory can be calculated by triangulation. The velocity of the meteor is measured by using a rotating shutter to interrupt the meteor trail up to sixty times per second. The velocity vector and the known position of the Earth in its orbit make it possible to compute meteor orbits.

The density of a meteoroid can be estimated from its deceleration in the atmosphere, showing that most meteoroids are of lower density than are meteorites. Statistical studies have shown that meteoroids with the greatest meteor heights have average relative densities of 0.6, while another group appearing about 10 kilometers lower have relative densities averaging 2.1. The few meteoroids that penetrate deep into the atmosphere have average relative densities of 3.7. Several hundred fireball meteors have been photographed, their masses ranging from 100 grams to 1,000 kilograms, including one meteorite fall near Lost City, Oklahoma. Experiments with artificial meteors and theories of meteor burning led to estimated initial meteoroid masses from observed optical effects. Meteors comparable in light to the brightest stars have initial masses of a few grams and diameters of about 1 centimeter, producing more than a megawatt of power. Meteor showers are produced by the most fragile (lowest-density) meteoroids, and different showers produce meteoroids of different character. In general, short-period comets exposed more often to the Sun produce higher-density particles than do long-period comets, because of greater evaporation.

During World War II, it was accidentally discovered that meteors could be detected by radar. The radar method of studying meteors is

The Geminid meteor shower over the mountains of Georgia in 1985. Unlike many meteor showers, the Geminids originate from the debris of an asteroid, 3200 Phaethon. (Jimmy Westlake) especially valuable because it can detect meteors in daylight and is sensitive to meteoroids as small as a microgram. This method depends on the fact that meteors separate electrons from atoms, producing ionized gases that can reflect radio waves. Meteor heights can be measured from the time delay of the return signal, and velocities can be determined from the frequency shift (Doppler effect). Observations from three stations are needed to calculate a meteoroid orbit. Several important meteor showers that occur only in daylight hours were discovered by radar, including the Beta Taurids, which are probably associated with Comet Encke. Radar also shows that radiants can be complex structures that appear to overlap and shift positions within a few hours.

Micrometeoroids with masses of a few micrograms or less have been collected by high-altitude aircraft and rockets. Micrometeoroids are fluffy particles containing carbonaceous material different from normal meteorites but consistent with comet theories. They can be studied with microphone detectors in space probes by measuring the intensity of their collisions. The weak structure of these particles indicates that they are gently separated from their parent material, suggesting dust emitted from evaporating ice in a comet, rather than violently ejected from high-temperature or colliding meteoroids. Particles of less than a milligram contribute the largest fraction of the total mass swept up by the Earth each day. Rocket-borne mass spectrometers have recorded metallic ions (charged atoms) of apparent meteoric origin, and meteor spectroscopy has provided chemical analysis of all the major meteor streams. These data indicate significant differences betwen cometary meteor material and the composition of meteorites.

Radioactive-dating techniques indicate that most meteorites have existed as solid bodies for about 4.5-4.7 billion years, close to the estimates for the ages of the Earth, Sun, and Moon. This suggests that all the matter of the solar system condensed at approximately the same time. Cosmic-ray dating from the amount of unusual isotopes produced in a meteorite by cosmic rays colliding with atoms in its crystalline structure usually indicates only a few million years since its formation, presumably by some fragmentation process from a larger asteroid. Fine bands are also observed in such meteorites, similar to those that occur in metal crystals subjected to sharp collisional shock. This finding has led to the idea that meteorites probably come from asteroids that were shattered in collisions.

CONTEXT

Meteors and meteor showers not only are interesting as visual phenomena but also provide one of the most important sources of information about asteroid and comet composition and deterioration, as well as clues to the origin of the solar system. Fortunately, most meteors are caused by very small particles (less than 1 gram) and are completely vaporized high in the atmosphere. Meteors enter the Earth's atmosphere with solar-system speeds and random inclinations to the ecliptic (plane of the Earth's orbit); thus, it appears that most meteors are associated with comets or with asteroids that have small inclinations.

A cometary origin for most meteors is supported by the phenomenon of meteor showers, which can be traced to particle swarms in various orbits with random inclinations around the Sun. Many of these showers can be associated with comets or former comets. They appear to be caused by particles released when cometary ices were evaporated by solar radiation. These particles either concentrate in a swarm of meteoroids behind the cometary nucleus or eventually become distributed in a stream around the entire orbit of the comet. Annual meteor showers occur when the Earth crosses a meteoroid stream, while more intense periodic showers occur when the Earth passes through a meteoroid swarm. The densities and compositions of these meteoroids are also consistent with a cometary origin.

The few meteoroids large enough to survive their passage through the atmosphere and yield meteorites are associated with the rare fireball meteors. Their trajectories tend to have low inclinations to the ecliptic, similar to asteroids. The crystalline structure of the metal in meteorites indicates that most were formed at high temperatures and slowly cooled over several million years. Thus, it appears that they did not come from icy comets; rather, they probably originated with asteroids. Calculations show that the rocky outer shell of an asteroid would insulate its hot metallic core, causing it to cool at the very slow rate suggested by the crystalline structure of iron meteorites. Furthermore, the cooling rate in a planet is too slow to fit this observed crystal pattern.

Meteoroids in asteroidal orbits enter the atmosphere with an average velocity of about 20 kilometers per second. Most are slowed rapidly by the atmosphere. If they survive as a meteorite, they simply fall to the ground at free-fall speeds and cool rapidly, since most of the hot surface material is swept off. Meteoroids larger than about one million kilograms (10-meter-sized) strike the ground with most of their initial velocities, producing impact craters. A 50-meter, 100-million-kilogram object moving at high speed can produce a one-kilometer-wide crater, causing widespread devastation by its shock waves and by throwing dust into the upper atmosphere, with marked effects on climate and life on Earth. Some evidence from large craters and geological layers of meteorite debris suggests the possibility that kilometer-sized objects strike the Earth about every twenty-six million years, coinciding with major extinctions of life-forms. One attempt to explain these data theorizes that the Sun has a dim companion star in a twenty-six-million-year eccentric orbit. At its closest approach to the Sun, its gravity would disturb many comets in the outer solar system, causing some of them to strike the Earth.

Joseph L. Spradley

FURTHER READING

Brown, Peter L. *Comets, Meteorites, and Men*. New York: Taplinger, 1975. This classic book is an excellent and very readable historical study of comets, meteors, and meteorites. Six chapters on meteor showers and meteorites cover the history of these phenomena from earliest times. Appendixes include a table of meteorite impact sites and the major annual meteor showers.

Burke, John G. *Cosmic Debris*. Berkeley: University of California Press, 1986. This book is a scholarly and well-documented history of meteorite and meteor discoveries and theories. Much of the book is suitable for the general reader, with interesting illustrations, but some parts are more detailed and technical. A 50-page bibliography contains about a thousand historical references.

Delsemme, A. H., ed. *Comets, Asteroids, Meteorites*. Toledo, Ohio: University of Toledo Press, 1977. This book is the result of an International Colloquium on the interrelations, evolution, and origins of comets, asteroids, and meteorites. Of the seventy-five articles, twenty-two are on meteors, meteoroids, and meteorites. Although the level is quite technical, the discussion offers a detailed firsthand account of research results for interested students.

Erickson, Jon. *Asteroids, Comets, and Meteorites: Cosmic Invaders of the Earth*. New York: Facts On File, 2003. Part of The Living Earth paperback series, a good basic resource for nonspecialists.

Levy, David H. *David Levy's Guide to Observing Meteor Showers*. Cambridge, England: Cambridge University Press, 2007. A thorough examination of different types of meteorites and a guide to observing meteors and meteor showers. For the general reader and the dedicated backyard observer.

_______. *The Quest for Comets: An Explosive Trail of Beauty and Danger*. New York: Plenum Press, 1994. Written by one of the codiscoverers of Comet Shoemaker-Levy 9, this book is for the general reader. It highlights the author's comet discovery program and the comet catastrophe theory.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. Geared for the undergraduate college student, this textbook treats planetary atmospheres as important physical features of the various members of the Sun's family. They are discussed both individually in the context of what is known about each planet's characteristics and with regard to theories about their evolution and the evolution of the entire solar system.

Time-Life Books. *Comets, Asteroids, and Meteorites*. New York: Time Life Education, 1992.

Part of the Voyage Through the Universe series. Heavily illustrated and for the general reader.

See also: Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Interplanetary Environment; Kuiper Belt; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Oort Cloud; Pluto and Charon.

Milky Way

Category: The Cosmological Context

Our Sun and solar system belong to the Milky Way galaxy, a large system of several hundred billion stars, gas, and dust. Beyond our galaxy are billions of other galaxies. The study of our galaxy adds to our knowledge of galaxies in general, while the study of other galaxies has revealed features later discovered in the Milky Way.

OVERVIEW

The Milky Way galaxy is a vast grouping of several hundred billion stars, gas, and dust to which the Sun and solar system belong. All the stars visible to the unaided eye in the night sky are a part of this same huge collection. Most of the stars, gas, and dust observed in our galaxy are contained in a thin disk more than 100,000 light-years in diameter, with a lens-shaped bulge at its center. This central bulge, also called the galactic nucleus, contains the greatest concentration of stars and, right at its center, a supermassive black hole. Within the disk are several spiral arms in which new stars form. Surrounding the disk and central bulge of our galaxy is the roughly spherical galactic halo, a few hundred thousand light-years in diameter, sparsely populated with old stars and old star clusters, but with virtually no gas or dust. The Sun and solar system are located in the disk, within one of the spiral arms or arm branches, about 26,000 light-years from the center. It takes the Sun and solar system more than 200 million years to complete one orbit around the galactic center.

This current picture of our galaxy had its origins in the 1600's, and it developed in its modern form during the twentieth century. However, speculations and myths about the Milky Way date back to ancient times, since it is such an obvious feature of the dark night sky. On any clear, moonless night away from city lights, the Milky Way is easily seen as a beautiful, pale white, diffuse band of light stretching across the sky. In ancient times it was given the name "Milky Way" because of its pale, milky appearance. (The English word "galaxy" comes from the Greek word *gala*, meaning "milk.") In some mythologies (including Greco-Roman), the Milky Way was the milk of the gods spilled on the sky. Some other cultures viewed it as the path by which departed souls ascended to the realm of the gods. In all, there were many imaginative mythological explanations for it.

In 1610, Galileo scanned the Milky Way with his small telescope and found it to be composed of lots of faint stars. By the 1700's, the arrangement of stars making up the Milky Way was described using terms such as "sheet," "disk," "millwheel," and "grindstone." In 1784, William Herschel developed a model of the Milky Way using a technique that he called "star gauging" in which he counted the number of stars he could see in his telescope's field of view when he pointed it in different directions. Based on his star counts, he described the Milky Way as a layer of stars with crooked or jagged outer edges and the Sun near its center. During the following century, Herschel's technique and model were refined, and by the early 1900's the Milky Way galaxy was pictured as an oblate ellipsoid, about 60,000 light-years in diameter and 13,000 light-years from top to bottom, with the Sun not far from the center.

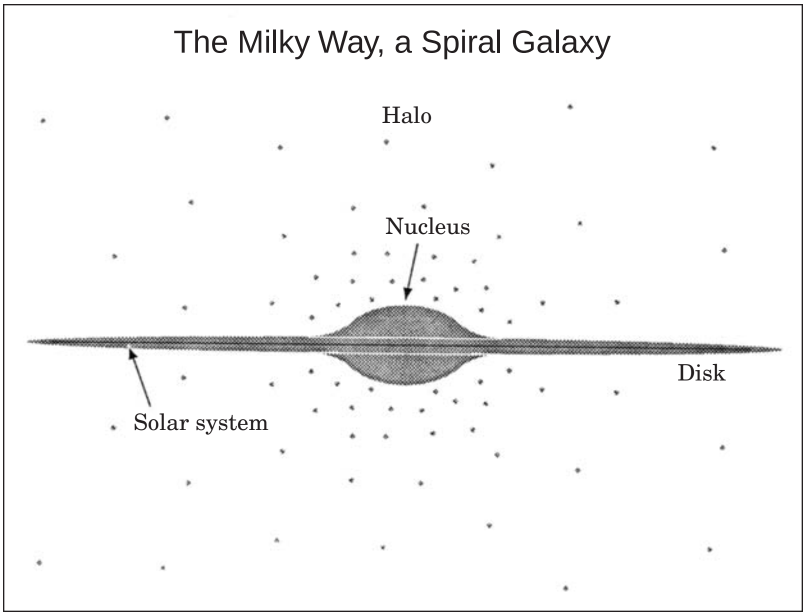
In 1918, Harlow Shapley of Harvard University concluded that the Galaxy was much bigger than previously thought, and the Sun was not near its center. Based on his study of the distribution in space of globular clusters (large clusters of about 100,000 to 1,000,000 very old stars), he determined that the galactic center was located about 50,000 light-years from the

Sun and solar system in the direction of the constellation Sagittarius. During the 1950's, the Milky Way's disk was found to have spiral structure. First, in the early 1950's, several spiral arm segments were traced optically using the locations of gaseous nebulae and young stars, objects that are concentrated in spiral arms. In the late 1950's, the first radio maps of the Galaxy's disk were produced, using radio radiation with a wavelength of 21 centimeters (and frequency of 1,420 megahertz), produced by hyperfine "spin-flip" transitions in neutral hydrogen atoms. The 21-centimeter radio maps traced the distribution of neutral hydrogen, which is concentrated in the spiral arms.

During the last half of the twentieth century, the Milky Way was observed over the entire electromagnetic spectrum, from radio waves to gamma rays, and this resulted in further refinements to our ideas about our galaxy. In current models, the galactic disk has an overall diameter of 100,000 to 130,000 light-years, possibly with a warped outer edge. The Milky Way seems to be a multiarmed spiral, with at least four and possibly more spiral arms, and smaller branches and spurs. There may be a central bar running through the galactic nucleus, with the spiral arms originating from the ends of it. Although visible light from the galactic nucleus is blocked by interstellar dust in the galactic disk between our location and the galactic center, the central region has been studied with radio waves, infrared, and gamma rays, all of which can penetrate the dust. The galactic nucleus is a packed concentration of stars, gas, and dust, with a supermassive black hole of several million solar masses at the center. The galactic nucleus and flat galactic disk are contained within a nearly spherical galactic halo of widely separated old stars and globular star clusters. An even more spread-out corona or outer halo extends beyond it, as far out as 150,000 to 300,000 light-years. The total mass of the Galaxy is calculated to be between several hundred billion and 2 trillion solar masses, based on the observed motions of stars and gas within our galaxy and small satellite galaxies on the fringes of the Galaxy.

However, only a fraction of the mass has actually been observed; the mass that has not yet been observed is called "dark matter," but astronomers and cosmologists are not sure what it is. The distance of the Sun and solar system from the galactic center has been reduced to half of Shapley's original value, about 26,000 light-years. The Sun, as well as most of the stars, gas, and dust in the galactic disk, has a nearly circular orbit around the galactic center. The orbital speeds of these solar-system bodies are between about 200 and 250 kilometers per second, and they take between about 200 million and 250 million years to complete one orbit around the galactic center. Some stars, particularly those in the galactic halo, have much more elongated orbits that carry them close to the center and then far out into the distant parts of the Galaxy.

The Cosmic Background Explorer (COBE), in 1990, captured this infrared image of the Milky Way, edge-on. (NASA)



Milky Way's interstellar gas, these clouds provide the cold, dense environment needed for the birth of new stars. Giant molecular clouds are concentrated around the central bulge and in the spiral arms.

The dust in our galaxy is confined for the most part to the central plane of the galactic disk. There it is very effective in obscuring the view of more distant objects at visible and shorter wavelengths. The Great Rift, the Coalsack, and the Horsehead nebula are some prominent dark concentrations of dust. However, longer wavelengths are able to penetrate the dust, and the longer the wavelength, the less the dust blocks the view. This provides a clue to the size of the dust grains. Small, solid particles are very effective at scattering electromagnetic radiation with wavelengths comparable to or shorter than the size of the grains. Therefore the size of the grains is a bit larger than the longest wavelengths of visible light, or about 0.001 millimeter. Polarization studies suggest the grains probably are elongated, perhaps needle-like. The dust grains probably are composed of iron alloys, silicate minerals, carbon compounds, various ices, or a combination of these.

Our galaxy appears to have at least four spiral arms, marked by clouds of gas and dust, as well as young stars that have recently formed in them. Objects near the galactic center revolve around it in less time than objects farther out. If spiral arms were permanent collections of objects, they would quickly "wind up" because the orbital period around the galactic center increases as the distance from the center increases. A solution to this "winding dilemma" is provided by the spiral density wave theory, which postulates that the spiral arms are produced by an underlying density wave pattern that propagates through the gas and dust of the galactic disk. The entire spiral pattern rotates with the same period around the galactic center.

KNOWLEDGE GAINED

Hydrogen, the most abundant element, exists in our galaxy as ionized and neutral single atoms, as well as diatomic molecules. Ionized atomic hydrogen is the most obvious form, because it can be seen at visible wavelengths as glowing clouds called H II regions. Hydrogen atoms are ionized by ultraviolet radiation from nearby hot stars, and visible light is produced when electrons recombine with the bare hydrogen nuclei. Clouds of neutral atomic hydrogen, called H I regions, were first detected in our galaxy in 1951, at the radio wavelength of 21 centimeters (frequency of 1,420 megahertz) emitted by hyperfine ("spin-flip") transitions in neutral hydrogen atoms. Astronomers had thought molecular hydrogen could not exist in interstellar space because ultraviolet radiation from hot stars would break the molecules apart into individual atoms. The discovery of interstellar molecules of carbon monoxide in 1970 spurred the search for other molecules. Since then, more than one hundred different molecules, including molecular hydrogen, have been detected in large, dense, cold molecular clouds, some of them hundreds of light-years across. The molecules in these clouds withstand the effects of ultraviolet starlight because they are shielded by interstellar dust. Making up about half of the

The stars, gas, and dust in the galactic disk move faster than the pattern, regularly overtaking the spiral-shaped density waves. Stars are not affected much by the slight increase in density, so they pass right through the density waves. Clouds of gas and dust, however, are slowed by the increase in density, so they pile up and are compressed, triggering star formation. The spiral arms are highlighted by young stars that have just recently formed and the gas clouds in which stars can form.

An intriguing feature of the Milky Way galaxy is its central region, since it is hidden from visual observations by thick clouds of dust. Radio radiation from the galactic center was first detected in 1932 by Karl Jansky, an engineer at Bell Telephone Laboratories. He was studying long-range radio communication for Bell Labs when he discovered steady radio static emanating from the galactic center in the direction of the constellation of Sagittarius. Obscured by interstellar dust at visible wavelengths, the galactic center received little attention until the 1960's and 1970's, when advances in radio and infrared astronomy made it possible to study.

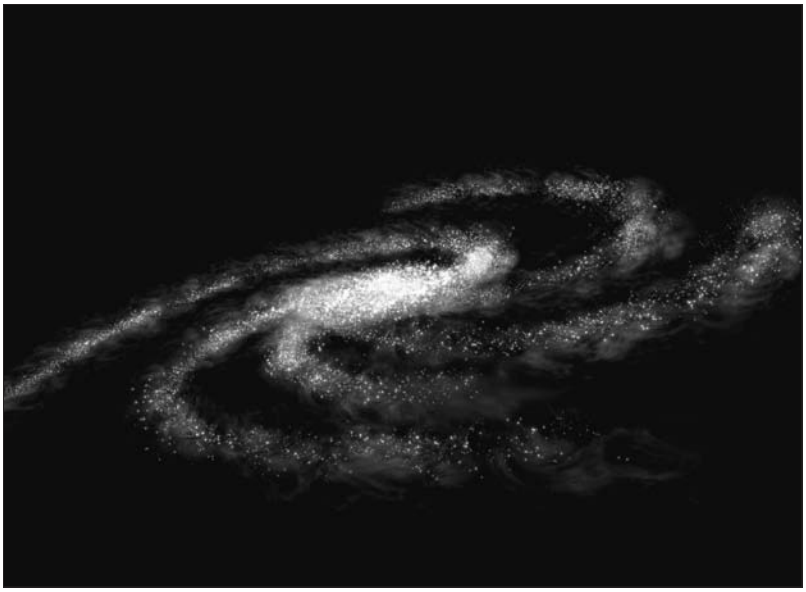
Infrared images of the central region revealed giant molecular clouds and lots of stars. Near the center, the average distance between stars is on the order of a few light-weeks (compared to a few light-years in the neighborhood around the Sun). A number of radio sources have been found in the central region, including several supernova remnants. The most prominent radio emitter there is an extended region named Sagittarius A (Sgr A). At the center of Sgr A, there is an extremely strong, very compact radio source named Sgr A* (pronounced "Sagittarius A star"). The orbital motion of several individual stars around Sgr A* has been recorded with high-resolution infrared imaging. These stars are at distances of only a few hundred to a few thousand astronomical units from Sgr A* and are moving at speeds of a few hundred to a few thousand kilometers per second, which indicates they are orbiting an object with several million solar masses. The only thing this massive that could fit inside the orbits of these stars is a supermassive black hole, which presumably is powering the Sgr A* radio source.

By the late 1970's, objects in the outer parts of the galactic disk were discovered to be revolving around the Galaxy at about the same speed as objects in the inner part of the disk, implying the existence of a substantial mass distributed throughout the Galaxy out to large distances. However, the mass that actually can be observed using any part of the electromagnetic spectrum falls far short of accounting for the needed mass. Originally referred to as "missing mass," the unobserved mass now is called "dark matter," since it is not really missing (at least if gravity behaves as physicists think it does). The most likely location for the dark matter is the galactic halo, which could contain considerable mass spread very thinly throughout its immense volume. Astronomers have yet to determine what dark matter is, but possible candidates include faint low-mass stars, dead stars, planets, rocks, black holes, neutrinos with mass, weakly interacting massive particles (WIMPs), massive compact halo objects (MACHOs), or some other form of matter.

CONTEXT

The Milky Way galaxy is only one of billions of observable galaxies that have a variety of sizes and shapes. The Milky Way is a reasonably large spiral galaxy. Galaxies also may have elliptical or irregular shapes, and they range in size from giant galaxies with at least several trillion solar masses down to dwarf galaxies with as little as several million solar masses. The distribution of galaxies in space shows a hierarchy. Galaxies tend to occur in groups called clusters of galaxies. In turn, clusters of galaxies tend to form larger groups called superclusters. The Milky Way galaxy is one of the two largest galaxies in a small cluster called the Local Group, consisting of about forty known member galaxies spread out over about 3 million light-years. The Local Group is part of a supercluster centered on the rich Virgo cluster of galaxies.

The continuing study of the Milky Way has advanced our knowledge of galaxies in general. A case in point concerns the origin and development of galaxies. Galaxies, including the Milky Way, are thought to have condensed from large clouds of gas, probably starting between one hundred million and one billion years after the



An artist's rendition of the spiral-armed Milky Way galaxy, in which our solar system is only a minute dot, in a universe composed of billions of galaxies. (NASA/JPL)

creation of the universe in the big bang about 13 to 14 billion years ago. This is confirmed by the ages of the oldest surviving stars in our galaxy, which are about 12 billion years old. Each protogalactic cloud was composed of hydrogen and helium, since those were the only elements produced in any significant amounts by the big bang. As the protogalactic cloud for the Milky Way contracted gravitationally, some parts were denser than others, and in these denser regions the first stars and star clusters formed. Since these stars and star clusters formed from material falling inward toward the center of the cloud, they developed elongated orbits that would carry them in close to the center and then back out into the outer fringes, where they spend most of the time. This became the halo of our galaxy. As the contraction proceeded, the protogalactic cloud began to spin faster to conserve angular momentum. The rapid rotation caused the protogalactic cloud to flatten, forming an equatorial disk around a central bulge of material. Thus the galactic disk and nucleus were born. The supermassive black hole at the galactic center probably developed very early because of the growing density as the galactic bulge or nucleus contracted. Massive stars synthesized chemical elements heavier than helium as they generated energy through nuclear fusion reactions in their interiors. When the massive stars exploded as supernovae, they dispersed the heavier elements and enriched the gas from which new generations of stars and planets would form. The Sun and solar system formed about 4.5 billion years ago, at which time about 2 percent by mass of the gas cloud from which it formed consisted of elements heavier than helium. Star formation continues to the present time in the disk of our galaxy, as it does in other spiral galaxies.

A variation of this basic scenario is that galaxies formed not from a single protogalactic cloud for each galaxy, but rather through the merger long ago of many smaller "seed galaxies." According to this version, small seed galaxies formed first after the big bang. After a first generation of stars formed in them, they began to merge, forming larger galaxies. There is growing evidence that our galaxy has "devoured" other smaller galaxies through a process known as galactic cannibalism, and is continuing to do so at the present time. Other galaxies show evidence of both galactic merger (two comparable size galaxies joining together) and galactic cannibalism (a larger galaxy devouring a smaller one).

Clark G. Reynolds and Richard R. Erickson

FURTHER READING

Bertin, G., and C. C. Lin. *Spiral Structure in Galaxies: A Density Wave Theory*. Cambridge, Mass.: MIT Press, 1996. Covers thirty years of spiral structure theory, separated into three sections: physical concepts, observations, and dynamical mechanisms. Can be technical.

Bok, Bart J., and Priscilla F. Bok. *The Milky*

Way. 5th ed. Cambridge, Mass.: Harvard University Press, 1981. The best nontechnical introduction to the Milky Way galaxy. Traces the history of observational discoveries and the development of models of the Galaxy based on the discoveries.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Devotes an entire chapter to the Milky Way galaxy.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Has an entire chapter on the Milky Way galaxy.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A thorough and well-written introductory college-level astronomy textbook. One chapter addresses the Milky Way galaxy.

Gingerich, Owen. "The Discovery of the Milky Way's Spiral Arms." *Sky and Telescope* 68 (July, 1984): 10-12. Based on interviews and technical publications, this excellent account by the leading historian of astronomy traces the changing model of the Milky Way from the 1920's through the 1950's, including the pioneering groundwork laid by Bart J. Bok and Walter Baade.

Herbst, William. "The Local System of Stars." *Sky and Telescope* 63 (June, 1982): 574-577. A somewhat technical historical and analytical survey of the problems involved in determining the Sun's actual location in the Galaxy, with special attention to Bok's contributions.

Tayler, Roger J. *Galaxies: Structure and Evolution*. Rev. ed. New York: Cambridge University Press, 1993. An easy to understand introductory book suitable for all readers.

Whitney, Charles. *The Discovery of Our Galaxy*. New York: Alfred A. Knopf, 1971. The best single-volume historical introduction to the attempt to investigate the nature of the Milky Way—from the ancients to 1970. Excellent illustrations.

See also: Big Bang; Comets; Cosmic Rays; Cosmology; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; General Relativity; Infrared Astronomy; Interstellar Clouds and the Interstellar Medium; Novae, Bursters, and X-Ray Sources; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Universe: Evolution; Universe: Expansion; Universe: Structure.

Miranda

Categories: Natural Planetary Satellites; The Uranian System

Miranda is the one satellite of Uranus about which there is good information. The whole Uranus system is different because Uranus is tilted. Miranda is different from any of the other natural satellites that have been studied.

OVERVIEW

Miranda is the eleventh of the twenty-seven known moons of Uranus, a "gas giant" planet that is the seventh planet from the Sun. Miranda was discovered by a Dutch-born American astronomer, Gerald Peter Kuiper, in 1948. Miranda was the daughter of the magician Prospero in William Shakespeare's play *The Tempest*. Features and places on Miranda are named for characters in Shakespeare's plays. When Miranda was discovered, photographic emulsions and techniques had just developed to the stage to be very useful in astronomy. Kuiper was using the McDonald Observatory to photograph the four satellites of Uranus that were known at that time: Oberon, Titania, Umbriel, and Ariel. Upon developing the film, he found a bright spot. A few days later, it was proven that the spot could not be a star and determined that it was, rather, a satellite of Uranus.

Miranda is synchronous with Uranus; that is, Miranda presents the same face to Uranus all the time. As a synchronous satellite, Miranda revolves about Uranus in its orbit in the same time that it rotates about its own axis:

33.9 hours (1.41 days). Miranda turns in the same direction as Uranus; that is, it has a prograde rotation. Miranda's orbit is about four degrees out of the plane of Uranus's equator; that is, its inclination is 4.3°. The inclination is large enough to cause doubt about whether Miranda was formed by the same process by which Uranus was formed, or if Miranda instead was captured in a close encounter.

Because Uranus is tipped on its side, Miranda's orbit is almost perpendicular to the orbit of Uranus. Miranda's eccentricity is small, meaning that the orbit is close to circular. The larger the eccentricity, the more elliptical an orbit; an eccentricity of zero means a circular orbit. Miranda's orbit is only slightly elliptical, with the maximum distance from Uranus being 132,000 kilometers and the minimum distance 128,000 kilometers. Miranda is a triaxial ellipsoid, with a 480-kilometer diameter along the axis pointed at Uranus. The equatorial axis is 468 kilometers, and the pole-to-pole axis is 466 kilometers.

The surface of Miranda is so fractured that it seems as if Miranda was torn apart and the pieces put back together in a haphazard manner. Some scientists have stated that Miranda was torn apart and reformed at least five times. Other ideas are that tidal forces caused partial differentiation. Differentiation is a separation of materials, with heavier rock sinking and forcing water ice to the surface. Near-infrared spectrometry has confirmed that the surface is ice. The surface shows it to be intensely cratered from meteor bombardment, although some of the craters have been reduced in height either by a change in the surface or by material ejected by other meteors covering up the crater as the material fell back to the surface. One crater is 25 kilometers across. There are more craters on Miranda than on other outer moons of Uranus. That is to be expected, since the closer a satellite is to the planet and its gravitational pull, the greater the density of meteors it encounters.

There are also coronae—large, 400-kilometer-wide areas of alternating light and dark stripes that are unlike anything seen in the universe. They look like racetracks with several tracks side by side. There are three of these on Miranda, named Arden, Inverness, and Elsinore. From space, Inverness looks like a chevron of bright material on a dark surface. The edges of the Arden and Inverness coronae are a trench with a cliff surrounding the coronae. These cliffs can be 10 to 20 kilometers high. The coronae seem to be at a lower altitude than the surrounding surface. Two of the coronae, Arden and Inverness, have albedos (reflectivity) that differ from that of the surrounding material. Elsinore does not have the exterior trench and has much the same albedo as the surrounding material. The coronae were formed in order: first Arden, then Inverness, and then Elsinore. Arden is in the leading hemisphere. Elsinore is in the trailing hemisphere. The sharp tip of the chevron edge of the Inverness corona is very close to the south pole of Miranda.

There are fracture lines and gorges running across the older cratered surface. There are few craters in the coronae, suggesting that they are not as old as the cratered surface area. Some scientists draw a correlation between the partial differentiation and the coronae. The coronae are a feature of partial differentiation, caused by a Miranda-Umbriel-Ariel orbital resonance. For some reason, the differentiation did not finish, but the rise of ice and sinking of heavier material formed the coronae.

Two other satellites of Uranus, Umbriel and Ariel, seem to interact with Miranda. A gravitational pull occurs when the satellites are close and then dissipates when they are far apart. This pull, along with the gravitational pull of Uranus, causes the interior of Miranda to flex. This tidal flexing may be the cause of the heat leading to the coronae and surface fractures. The gravitational interaction also causes a change in the orbit of Miranda.

A surface temperature of 86 kelvins, determined by Voyager 2, may be enough to melt an ammonia-water mixture. Near-infrared spectrometry not only shows that the surface contains a large amount of water ice; there is also a feature in the spectra that may indicate the presence of ammonia hydrate. This mixture would lower the melting point enough that tidal flexing could account for the fracturing and coronae. The other surface material is carbonaceous and/or silicate matter. The density of Miranda (1.25 grams/centimeter^3) is close to that of ice (1.0 grams/centimeter^3) so there cannot be a large percentage of other material. The rock fraction can only be 33 percent.

The albedo of Miranda was measured by Voyager at 32 percent. Hubble Space Telescope (HST) measured it at a lower value. One question is, therefore, whether the change in value indicates a real difference, meaning that the northern hemisphere imaged by HST is darker than the southern hemisphere imaged by Voyager, or whether there is another explanation for the difference in values.

There are at least four theories about how Miranda and the other satellites were formed. One is the co-accretion model. Solid particles that were a part of the solar nebula were collected by Uranus's gravitational force to form a disk that later coalesced into satellites. The accretion disk model, by contrast, postulates that the gravitational force of Uranus collected material from the solar nebula. This material formed a disk and later collected into a satellite. A third model is the spin-out model, which suggests that some of the material spun out from the planet as it contracted, forming Miranda and the other satellites. The blow-out model theorizes that an Earth-size body struck Uranus, causing Uranus to tilt and ejecting material that formed a disk and later satellites. Some also subscribe to the theory that Miranda was a small body captured by Uranus's gravitational force as it came close to the planet.

KNOWLEDGE GAINED

Miranda can be seen with Earth-based instruments, but it is so close to Uranus that it is difficult to study its characteristics. Near-infrared spectra obtained using the United Kingdom Infrared Telescope (UKIFT) revealed much of what we know about the surface composition of Miranda. Earth-based instruments also serve well to study orbits, even for a faraway satellite like Miranda. Once Miranda was known, it was clear that it had been seen in other photographs. Putting all of the sightings together produced a knowledge of the satellite's orbit.

Goals for the Voyager studies for Miranda were twofold: observational and magnetic. The magnetic goals were to see if Miranda had a magnetic field (it was determined that it does not), how Miranda interacted with Uranus's field, and if there were charged particles in the atmosphere. The observational goals for Voyager were to note the satellite's size, shape, mass, density, shape changes with time; types of surface structures; surface processes; and albedo change with phase angle. Voyager data concerning albedo are similar but slightly lower than that obtained from the Hubble Space Telescope. Voyager data and years of observations from Earth were used together to calculate better values for Miranda's orbit, eccentricity, and angle of inclination. The measurement of the mass of Miranda was made from radio science data. Imaging data were used to determine size, which is more complicated than merely measuring distances in a single image. Several images were needed to cover Miranda. By reference to the same feature in at least two images, researchers then generated a mathematical model to calculate where each feature should be. From that model the size could be calculated.

Comprehensive photometry has been done of Miranda using the HST. All possible phase angles from Earth were studied. A phase angle is the angle between the Sun-to-Miranda line and the Earth-to-Miranda line. The geometric albedo and opposition surge were measured for Miranda. Opposition surge is the brighter reflection that occurs at phase angle of zero.

Together, the data from Voyager and the HST not only vastly increased our knowledge of Miranda but also added to what is known of Uranus's satellites in general. Prior to these space-based observations, the number of known satellites of the gas giant was five; it is now known that there are at least twenty-seven Uranian satellites.

CONTEXT

Data gained by a study of Miranda may be second in importance to the new ideas generated by its unusual surface features. Scientists have to develop new explanations for the generation of the features of Miranda. Some information is known, but the total explanation will require innovative ideas.

Ideas about tectonics and how surfaces of worlds can move and change will help scientists to understand the creation of Earth and the separation of the continents. Studies of tidal flexing, such as that which occurs in the Uranian system, may be important to the study of Earth's plates tectonics. The effect of gravity on the tides is visible, but is that all the effect that gravitational pull has on the Earth? Scientists are also learning more about melting points and sublimation values of mixtures as they try to explain Miranda.

The best way to study Miranda will be to mount another interplanetary probe such as Voyager, but until such a mission can be planned, new techniques are being developed for using land-based instruments and the Hubble Space Telescope to study Miranda and other distant worlds. Humanity will have to improve the instrumentation used presently, and new devices and techniques will have to be designed and tested to expand the study of Miranda.

C. Alton Hassell

FURTHER READING

Bond, Peter. *Distant Worlds: Milestones in Planetary Exploration*. New York: Copernicus Books, 2007. The author discusses each system; its parts, such as planet, moons, and rings; and how each mission has developed our knowledge of that system. Illustrations, bibliography, appendix, index.

Corfield, Richard. *Lives of the Planets*. New York: Basic Books, 2007. The author takes the reader through the different space missions that have contributed to our knowledge of the planets. The book is divided by planets, and the information gathered by each mission is discussed. Index.

Croswell, Ken. *Ten Worlds: Everything That Orbits the Sun*. Honesdale, Pa.: Boyds Mills Press, 2007. Each system is discussed in turn. Written for a younger audience, this book provides good basic information. Illustrations, bibliography, index.

Elkins-Tanton, Linda T. *Uranus, Neptune, Pluto, and the Outer Solar System*. New York: Chelsea House, 2006. This book explores the Sun's relationship with the three outer planets and their moons. It looks at these planets as recorders of the formation of the solar system. Aimed at a general or high school audience. Illustrations, bibliography, index.

Loewen, Nancy. *The Sideways Planet: Uranus*. Mankato, Minn.: Picture Window Books, 2008. An educational children's book devoted to the planet Uranus. Covers the planet's rings, moons, and tilted axis.

McFadden, Lucy-Ann Adams, Paul Robest Weissman, and T. V. Johnson, eds. *Encyclopedia of the Solar System*. San Diego: Academic Press, 2007. The editors have collected articles written by many experts. It is one of the best surveys of material about the solar system. Illustrations, appendix, index.

Miller, Ron. *Uranus and Neptune*. Brookfield, Conn.: Twenty-First Century Books, 2003. Considers Uranus and its satellites in comparison with other gas giants, especially Neptune, including their atmospheres.

Miner, Ellis. *Uranus: The Planet, Rings, and Satellites*. New York: Ellis Horwood, 1990. The author thoroughly covers the topics of both the Uranian system and the Voyager mission. Miranda is featured as a remarkable satellite. Illustrations, bibliography, index.

See also: Auroras; Callisto; Enceladus; Eris and Dysnomia; Europa; Ganymede; Iapetus; Io; Jupiter's Satellites; Lunar Craters; Lunar History; Lunar Interior; Lunar Maria; Lunar Regolith Samples; Lunar Rocks; Lunar Surface Experiments; Mars's Satellites; Neptune's Satellites; Planetary Formation; Planetary Satellites; Pluto and Charon; Saturn's Satellites; Titan; Triton; Uranus's Satellites.

N

Nemesis and Planet X

Category: Planets and Planetology

Scientists developed the Nemesis and Planet X theories in an effort to explain why mass extinctions on Earth appear to have occurred roughly every 26 million years. Each theory suggests that a still undiscovered star (Nemesis) or planet (Planet X) periodically perturbs the orbits of comets, sending them into the inner solar system where some could collide with Earth.

OVERVIEW

Nemesis and Planet X are theoretical astronomical bodies whose existence has been suggested as a possible driving force for periodic mass-extinction episodes. Nemesis and Planet X theories both propose that about every 26 million years movement of an astronomical body through space causes massive extraterrestrial objects to collide with Earth. The resulting catastrophic changes in the Earth's environment led to the widespread extinction of species. The periodicity of the mass-extinction episodes, and the impacts causing them, are themselves theoretical.

The Nemesis theory hypothesizes that the Sun is orbited by a companion star so small, dim, and distant that astronomers have yet to discover it. This dark, low-mass star is thought to move in a highly elongated orbit at a distance between 25,000 to 150,000 astronomical units (AU) from the Sun. About every 26 million years, when it most closely approaches the Sun, the companion star is believed to pass through the huge "comet reservoir" known as the Oort Cloud. This vast region, which may contain anywhere from 100 billion to 10 trillion widely separated comets, surrounds the solar system, extending to a distance of perhaps 50,000 AU from the Sun. As the companion star moves through the Oort Cloud, its gravitational field disturbs the orbits of nearby comets. Some are deflected into deep space, while others are pushed toward the inner solar system. Of those comets entering the inner solar system, it is likely that some could collide with Earth. Because of the role that the Sun's companion star is thought to play in causing periodic devastating impacts on Earth, researchers have dubbed it Nemesis, after the Greek goddess of vengeance.

The Planet X theory, which also involves cometary perturbation, proposes that the solar system has a massive, undiscovered tenth planet whose orbit lies beyond that of Pluto. The "X" represents both "the unknown" and the Roman numeral X (ten). Researchers have suggested that this Planet X may have so far escaped detection because its orbit lies outside the plane of the solar system. Planet X is believed to be at least as large as Earth and 50 to 100 astronomical units from the Sun. It is thought to orbit the Sun in about 1,000 years, along a highly elongated, eccentric path. According to this theory, the orbit of Planet X slowly precesses over a period of 58 million years—that is, it regularly changes orientation in space. The precession is such that about every 28 million years (roughly one-half the precession period), Planet X perturbs comets within the Kuiper Belt. This disk of comets located beyond Neptune at a distance of roughly 35 to 1,000 AU from the Sun is regarded by some astronomers as merely the inner portion of the Oort Cloud. Although Planet X has cleared a path through the Kuiper Belt, roughly every 28 million years its precession brings it close to the margins of the cleared area, where comets fall under its gravitational influence. As in the Nemesis scenario, gravitational forces deflect some comets into the inner solar system, leading to periodic cometary impacts on Earth.

Many scientists link such cometary impacts to mass extinctions experienced throughout Earth's history. Earth's fossil record shows that over the past 600 million years there have been several mass-extinction events, forming the boundaries between the major divisions on the geologic timescale. For example, it was in the boundary between the Cretaceous and Tertiary periods (also known, from the German, as the K-T boundary) that the dinosaurs, nearly all marine plankton (tiny photosynthesizing plants), and 15 percent of marine invertebrate families became extinct. Evidence strongly suggests that the K-T extinction 65 million years ago was related to extraterrestrial impact.

If a comet or other large extraterrestrial body were to collide with Earth, the environmental effects would be severe and widespread. The immediate area of impact would be devastated by heat energy and a powerful shock wave; tsunamis and earthquakes might also result. Massive quantities of dust and gases would be thrown into the atmosphere, and a portion of the atmosphere itself might be ripped away. Fireballs formed by coalescing dust and gases would set off large-scale wildfires. The resulting smoke, combined with the dust and gases, would block sunlight for months, reducing global temperatures and impeding photosynthesis.

Depending on where the collision occurred, conditions could be even more severe. If the impact site were rich in limestone and evaporite minerals (as is believed to be the case for the K-T event), sulfur and water vapor released from the rock could combine with atmospheric gases to produce nitric and sulfuric acid, which would eventually fall to Earth as acid rain. In combination with atmospheric gases, carbon freed from the limestone would produce carbon dioxide, a "greenhouse gas" capable of trapping heat and causing global warming. Plant and animal species unable to withstand the extremes and fluctuations in temperature, disruptions in the food chain, and other inhospitable conditions would face extinction.

METHODS OF STUDY

The Nemesis and Planet X theories span several disciplines. Paleontologists, geologists, chemists, physicists, and astronomers are among the scientists whose studies have led to the development of these and associated concepts. Both theories are rooted in the idea that extraterrestrial impacts have caused or contributed to extinction events and that these impacts occur periodically.

Some researchers have used iridium data to link extraterrestrial impacts to mass extinctions. Iridium, an element related to platinum, is rare in the Earth's crustal rock (its average abundance is 0.001 part per billion). Gravity drew most of this extremely dense and unreactive element to the Earth's core when the planet was still molten. By contrast, laboratory analysis of meteorites (rocks of extraterrestrial origin found on Earth) has detected concentrations on the order of 500 parts per billion. The iridium found in surface sediments typically originates from space and is most often deposited as a light but steady rain of small cosmic particles. Where elevated iridium concentrations are present, extraterrestrial impact may have occurred. (Some scientists favor the alternative theory that iridium has reached the Earth's

Nemesis and Planet X theories both propose that about every 26 million years movement of an astronomical body through space causes massive extraterrestrial objects to collide with Earth (artist's rendition). (Don Davis, NASA) surface through volcanic activity.) High concentrations of iridium have been found at the Cretaceous-Tertiary boundary in rock and sediment samples from around the world. While iridium anomalies have also been noted in association with other boundaries marking major extinction events, the correlations are less striking and more problematic. To determine iridium concentrations in a sample, chemists use neutron activation analysis, whereby material to be analyzed is exposed to neutrons from a nuclear reactor. Once the iridium has been made radioactive, a gamma-ray detector system determines the element's concentration.

Between the Cretaceous and Tertiary periods, about 65 million years ago, a catastrophic event occurred, which many scientists believe to be an asteroid or cometary impact, that resulted in the extinction of nearly all marine plankton, 15 percent of marine invertebrate families, the large dinosaurs, and many other species. (Don Davis, NASA)

Researchers have identified other signs of possible extraterrestrial impact, all within Cretaceous-Tertiary boundary sediments. Grains of quartz and other minerals have undergone shock metamorphism, a change in crystal structure and density that results from the passage of high-pressure shock waves. Also present are tektites, glassy silicate-rich balls of varying composition, which are believed to form when rock and soil vaporized by an extraterrestrial impact recondense and fall back to Earth. Soot in boundary-layer samples from around the world suggests extensive burning. In addition, geophysical methods have revealed a huge, half-submerged, 65-million-year-old crater, Chicxulub, off Mexico's Yucatán Peninsula. An impact structure almost 180 kilometers in diameter, Chicxulub Crater is buried beneath roughly 2 kilometers of sediment and was detected during oil exploration by means of magnetism and gravity surveys.

To determine whether extraterrestrial impacts are periodic in nature, paleontologists have compiled vast amounts of data on the fossil record of life, notably data on when organisms appeared and disappeared. Researchers noted a possible periodicity in extinction events after creating graphs and conducting statistical analyses using computerized data indexes. Some have also performed similar computer analyses using a list of known impact craters and their ages; periodicities of 28 million to 31 million years have been reported. It should be noted, however, that there is considerable uncertainty in the dates assigned to many of the craters.

Scientific evidence suggesting impact-related extinction and periodic extinction episodes led scientists to develop the Nemesis and Planet X theories. Both theories assume that the extraterrestrial agents in question are comets. Asteroids, rocky astronomical bodies smaller than planets, can also strike the Earth and wreak havoc. However, their travels are considered too erratic to account for the periodicity discussed here. To validate either the Nemesis or Planet X theory, scientists must find an astronomical body whose observed characteristics correlate with predictions.

The Planet X theory is generally considered to be the weaker of these two proposed models. In 1985, when the theory was first published, the existence of a massive tenth planet provided a viable explanation for minor irregularities in the movements of the outer planets. (Pluto's mass proved insufficient to account for the anomalies.) Researchers suggested that a tenth planet had escaped detection because efforts to find it had concentrated on the plane of the solar system, while Planet X's orbit was inclined in relation to the plane. By the mid-1990's, these arguments were less convincing. Although astronomers have an idea where to look, they have yet to observe Planet X. In fact, the lack of orbital irregularity in the outer planets makes it unlikely that a planet massive enough to perturb comets lies in the area predicted by the Planet X theory.

The search for a star matching Nemesis's description continues. Because more than half the known stars are believed to have a companion, the existence of a companion to the Sun would not be unusual. If this companion were a red dwarf star, too small and too dim to be observed readily despite its relative proximity to the Sun, it could escape detection. Moreover, because astronomers have yet to determine the distances to many red dwarfs, it is possible that Nemesis has been observed but not recognized. At present, Nemesis would be halfway through its orbital cycle, near its greatest distance from the Sun (and roughly 13 million years away from its next impact-triggering passage through the Oort Cloud). Researchers are scanning the skies using a computer-controlled reflecting telescope to automatically record and compare images of red dwarf stars. This computerized system observes each candidate over time and uses parallax shift to assess how far away it is. Parallax is the apparent displacement of an object against a background of more distant objects when it is viewed from a different location; the movement of the Earth in its orbit provides the necessary change in viewpoint. The closer an object is to the viewer, the greater the parallax; because

Did an Asteroid Kill the Dinosaurs?

In 1977, Luis W. Alvarez and his son Walter began an investigation of a single-centimeter layer of clay that was sandwiched between two limestone strata containing large deposits of Cretaceous-Tertiary fossils. Such fossils were significantly absent elsewhere in the sample. The clay deposit dated from the boundary between the Cretaceous and Tertiary periods, referred to as the “K-T boundary,” roughly 65 million years ago, when the dinosaurs disappeared and modern flora, apes, and large mammals appeared.

Alvarez and his son used a trace of iridium in the composition of the clay sample to determine how long it had taken for the clay to be deposited and so to calculate the time that had elapsed during the Cretaceous-Tertiary transition. Iridium is basically an extraterrestrial substance. All the iridium in Earth's crust is only one ten-thousandth of the iridium abundant in meteorites. Alvarez selected iridium because it was the best material to use in determining the amount of debris that fell on Earth during this crucial period. Iridium is deposited uniformly around Earth, and Alvarez wanted to account for these uniform deposits. He began with the theories of Sir George Stokes, who formulated the viscosity law, a calculation of the rate at which small particles fall in the air. Stokes had based this law on his observations of the fallout of ash from the huge eruption of the Krakatoa volcano near Java in the 1880's. After discounting many possible hypotheses, such as a gigantic volcanic eruption, a supernova, or Earth's passing through a cosmic cloud of molecular hydrogen, Alvarez developed the hypothesis that an asteroid had collided with Earth.

According to his calculations, the asteroid had to be 10 kilometers in diameter. Its impact would have been catastrophic, far exceeding the worst nuclear scenario yet proposed. As Alvarez said,

> The worst nuclear scenario yet proposed considers all fifty thousand nuclear warheads in U.S. and Russian hands going off more or less at once. That would be a disaster four orders of magnitude less violent than the K-T asteroid impact.

Alvarez knew that the margin for error in discoveries was exponential, because of the possibility of mistakes in the data. As more data were collected from other sources, however, the argument only became stronger. Although the asteroid hypothesis has not been fully accepted by the scientific community, a number of predictions based on the theory have been verified experimentally and by computer simulation.

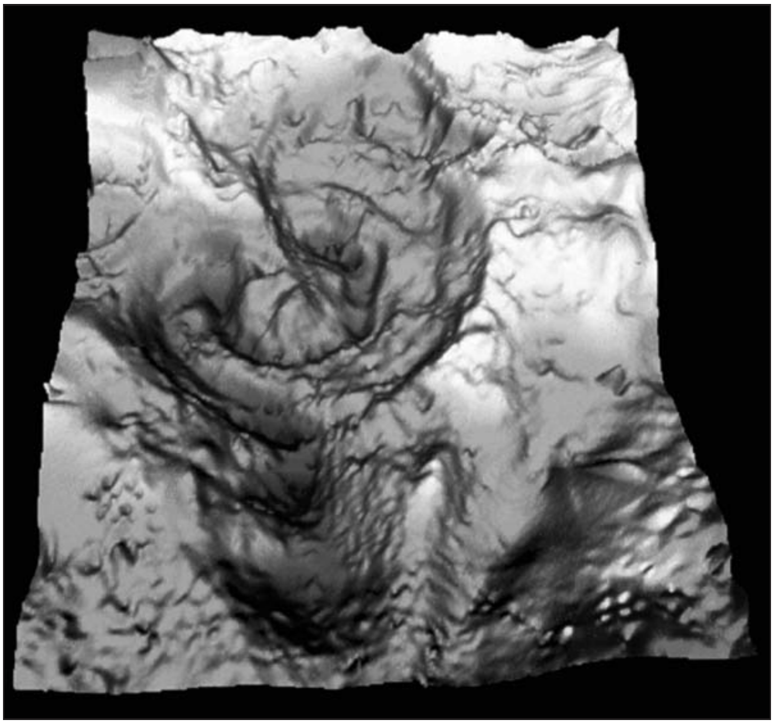
Nemesis should be the closest star to Earth other than the Sun, it should exhibit greater parallax than others. So far researchers surveying the northern hemisphere have eliminated more than half of the 3,100 red dwarf stars under scrutiny. The search will continue until astronomers have eliminated all candidates or until Nemesis is found.

CONTEXT

The Nemesis and Planet X theories both arose in response to other theories dealing with the nature of extinction on Earth. The idea that collision with an extraterrestrial object could trigger mass-extinction events had been considered before 1980; however, that year saw the first publication of compelling evidence supporting an impact theory. A multidisciplinary research team reported in the journal *Science* that high concentrations of iridium detected in rocks at the Cretaceous-Tertiary boundary were possible evidence that an extraterrestrial agent had ended the reign of the dinosaurs. Geologist Walter Alvarez, physicist Luis Alvarez, and nuclear chemists Frank Asaro and Helen V. Michel proposed that an asteroid or comet roughly 10 kilometers in diameter struck the Earth 65 million years ago, severely affecting the environment and depositing iridium. Their post-impact scenario, according to which the vast quantities of dust and gases flung into the air affected global photosynthesis and temperatures, caught the attention of researchers contemplating the effects of nuclear warfare. They applied the Alvarez team's projections in developing a scenario similar to the "nuclear winter" model advanced by Cornell astronomer Carl Sagan in the early 1980's.

In 1984, paleontologists David Raup and J. John Sepkoski, Jr. published their findings on the periodicity of extinctions. Using Sepkoski's extensive compendium of fossil-record data, the research team was able to distinguish some minor extinction episodes from background extinction levels. They found that over the past 225 million years significant extinction events had apparently occurred approximately every 26 million years.

The scientific community responded quickly with ideas that united and relied upon both the impact and periodic-extinction theories. In April, 1984, the journal *Nature* featured several articles exploring the astrophysical aspects of Raup and Sepkoski's 26-million-year periodicity. Among them was a paper by Walter Alvarez and physicist Richard Muller, who found a 28.4-million-year periodicity in large, well-dated impact craters. In a separate article that gave Nemesis its name, Muller, along with astronomers Marc Davis and Piet Hut, proposed the existence of a small, dim companion to the Sun, a "death star," that would periodically trigger comet showers on Earth.



The 65-million-year-old Chicxulub Crater, off Mexico's Yucatán Peninsula, is buried beneath 2 kilometers of sediment. Gravity measurements have resulted in this computer-generated gravity map of the crater. (Virgil L. Sharpton, University of Alaska, Fairbanks)

Astronomers Daniel Whitmire and Albert Jackson independently reached the conclusion that a distant solar companion drives the extinctions. In January, 1985, Whitmire and astronomer John Matese published a paper in *Nature* that presented Planet X as the driving force for periodic extinction events. Since then, scientists have continued to investigate the possible connection between extraterrestrial impact and extinction, studying the distribution and significance of elevated iridium concentrations, exploring the idea of periodic extinction, and searching for the mechanisms that may drive mass-extinction events.

Nemesis, Planet X, and related theories have generated considerable controversy and interest within the scientific community. Geologists and paleontologists learn that "the present is the key to the past" and that the natural, often gradual forces observable today can explain past geologic events. Hence, to many of these scientists the idea of sudden, violent, global catastrophe, the likes of which humankind has never experienced, seems more like science fiction than sound science. These theories also challenge the notion that evolution's primary driving force is an internal force—competition among species. If cometary impact does indeed occur every 26 million years, external forces such as environmental change may play a more significant role in evolution than scientists previously assumed.

The connection of an asteroid impact (at Chicxulub Crater) to the mass extinction at the Cretaceous-Tertiary boundary is well accepted within the scientific community. However, in the early twenty-first century, some researchers proposed an asteroid impact to explain the even greater mass extinction, often referred to as the Great Dying, that occurred 248 million years ago at the Permian-Triassic boundary. A crater much larger than Chicxulub, buried under 1.5 kilometers of ice in Antarctica, is dated to about 248 million years. That crater also is positioned such that the Siberian Traps are located at the antipode of this crater. This hypothesis suggests that seismic energy from the enormous impact in Antarctica underwent antipodal focusing through the Earth's core to devastate the area now known as the Siberian Traps. That area also suffered tremendous volcanic activity about 248 million years ago. It must be said that this scenario was quite controversial in 2008, but then so was the original Alvarez concept nearly three decades earlier. More research is required before the Permian-Triassic mass extinction in which over 95 percent of all life on Earth died out might be explained in this fashion.

Nemesis and Planet X theories lost favor with time. A number of researchers pointed out that a red dwarf far from the Sun in a rather unstable orbit would likely escape into interstellar space after only a few orbits. Continued observations failed to find a planet beyond Neptune large enough to account for the disruption of small bodies that then enter the inner solar system.

It must also be pointed out that in 2006 the International Astronomical Union approved a new classification system for bodies in the solar system. As a result, Pluto was demoted to the status of a dwarf planet. This left the solar system officially with eight full-fledged planets. Thus the "X" in the Planet X theory could still stand for unknown but would be incorrect if associated with a yet-to-be-discovered tenth planet beyond Pluto.

Karen N. Kähler

FURTHER READING

Close, Frank. *Apocalypse When? Cosmic Catastrophe and the Fate of the Universe*. New York: William Morrow, 1988. Chapter 4 discusses Nemesis and Planet X. Chapter 5 deals with the Cretaceous-Tertiary extinction and the Alvarez team's research. Includes a glossary. Written for a nontechnical audience.

Dauber, Philip M., and Richard A. Muller. *The Three Big Bangs*. Reading, Mass.: Addison-Wesley, 1996. Section I, which focuses on catastrophic impacts, devotes a chapter to Nemesis and related theories concerning mass extinction. An accessible, well-written book suitable for the general reader.

Goldsmith, Donald. *Nemesis*. New York: Walker, 1985. An accessible book exploring extraterrestrial impacts, the possibility of mass-extinction cycles, and the scientific controversy surrounding the Nemesis and Planet X theories. For the general reader.

Gould, Stephen Jay. *The Flamingo's Smile*. New York: W. W. Norton, 1985. Includes well-written essays addressing Nemesis and related theories (chapter 30), periodic extinction (chapters 15 and 28), and extraterrestrial impacts (chapter 29). For the general reader.

Gribbin, John, and Mary Gribbin. *Fire on Earth*. New York: St. Martin's Press, 1996. Chapter 8, which discusses Nemesis, Planet X, and the basis for periodic-impact theory, presents detailed arguments for and against the Nemesis theory. A thorough and accessible work intended for the general reader.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar system astronomy suitable for an introductory college course but also accessible to nonspecialists. Filled with supplemental learning aids and solved student exercises. A Web site is available for educator support.

Muller, Richard. *Nemesis*. New York: Weidenfeld & Nicolson, 1988. Written by a member of the research team that gave Nemesis its name, this book is an engaging firsthand account of the development of the Nemesis theory and its impact on the scientific community. For the general reader.

Raup, David. *The Nemesis Affair*. New York: W. W. Norton, 1986. An accessible account of the genesis of the Nemesis theory and its reception by the scientific community, by one of the chief proponents of the periodic extinction theory. Suitable for the general reader.

Reid, Neil, and Suzanne Hawley. *New Light on Dark Stars: Red Dwarfs, Low-Mass Stars, Brown Stars*. 2d ed. New York: Springer Praxis, 2005. A technical description of those stars that are not very luminous. Discusses discoveries of brown dwarfs and extrasolar planets.

See also: Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Impact Cratering; Kuiper Belt; Oort Cloud; Pluto and Charon.

Neptune's Atmosphere

Categories: The Neptunian System; Planets and Planetology

Neptune's atmosphere exhibits surprisingly rapid changes, large cloud systems, extremely fast winds, and large-scale vertical motions associated with heating from outside and inside. These features were quite unexpected for a planet so far away from the Sun, where astronomers expected to find a cold, featureless atmosphere.

OVERVIEW

Current models for Earth's atmosphere assume that the primary source of energy is solar radiation. Some basic features of a planetary atmosphere can be predicted from the planet's mass and surface temperature. The mass has to be large enough that the escape velocity of gas at the surface is higher than the random thermal motion speed of most of the molecules. The temperature must be neither too high (as at Mercury) nor so low that everything condenses and freezes. Atmospheric pressure at any level can then be related to the weight per unit area of the gas above that level.

Neptune's mass is 1.02×10^{26} kilograms, which is roughly 17.09 times that of Earth. It radiates nearly twice as much energy as it receives from the Sun, so that its internal heat is a major factor in the atmospheric dynamics.

The "surface" of Neptune is arbitrarily defined as being the level where the gas pressure reaches 1 bar or 101.325 newtons per square meter, this being the value of standard sea-level pressure on Earth. The planet's radius is defined at this level. At the equator, this radius is 24,764 kilometers. Here the acceleration due to gravity is 10.11 meters per second squared, close to the Earth surface value of 9.81 meters per second squared.

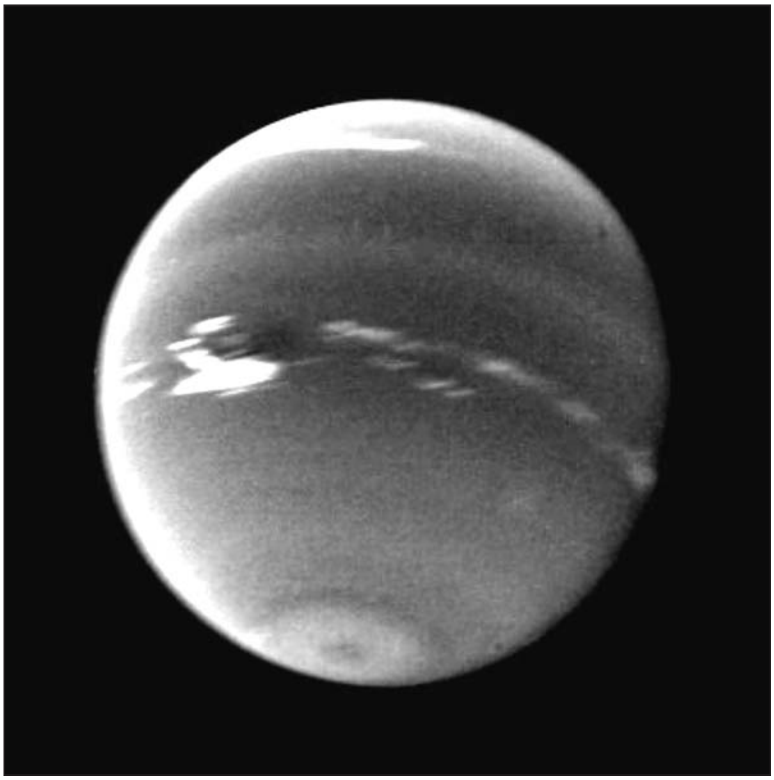
However, the gaseous atmosphere extends down to levels where the pressure exceeds 1,000 bars. The boundary between the condensed (solid and liquid) and gas-state substances is ill defined but is believed to occur at about the same radius as the Earth's surface radius, which is only about one-quarter of the radius of

Neptune. Substances such as hydrogen, helium, and methane exist as solids called "ices," even though their temperature is on the order of 700 kelvins or more.

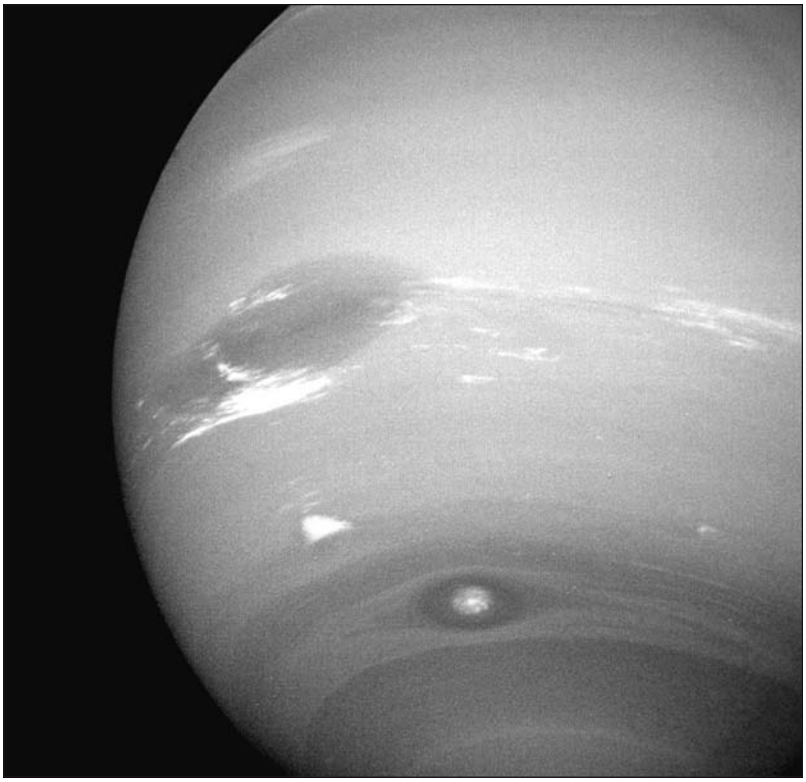
The temperature decreases outward from about 130 kelvins at 10 bars to 59 kelvins at 1 bar, the arbitrary gaseous "surface" of zero altitude. At altitudes where the pressure drops from about 3 bars to 1 bar, the temperature lapse rate on Uranus, which is similar in many respects to Neptune, is on the order of 1 kelvin per kilometer. The temperature drops to a minimum of around 50 kelvins at an altitude of roughly 50 kilometers, where the pressure is 0.1 bar. This is taken as the tropopause, or top of the troposphere. The temperature then rises almost linearly to about 150 kelvins at about 200 kilometers, where pressure is 0.5 millibar. Temperature remains constant at this level to about 500 kilometers, where pressure is 0.001 millibar, defining the stratosphere. It then rises again, reaching 275 kelvins at roughly 800 kilometers, where pressure is 1×10^{-5} millibar. This is the mesopause. At these high levels, the temperature is believed to be due to any of three mechanisms: the "dayglow," where diatomic hydrogen gets dissociated by collision with low-energy electrons; heating by production of auroras due to interaction of the magnetic field with ions; or joule heating as ions are accelerated by the magnetic field, relative to the movement of the neutral gases in atmospheric winds.

The average composition of Neptune's atmosphere is 77 to 83 percent hydrogen, 16 to 22 percent helium and 1 to 2 percent methane, with about 0.019 percent hydrogen deuteride and traces of ethane. Oxygen is present in water, nitrogen in ammonia, and carbon dominantly in methane, though some carbon monoxide has been detected on Neptune. Other elements, such as sulfur, have also been detected. It is believed that in the molten core of the planet there is silicate rock.

Condensation clouds are present at various levels in the Neptune atmosphere. Clouds of water, ammonia, and hydrogen sulfide are predicted to form with a base at a level where pressure is more than 500 bars and the temperature around 480 kelvins. Around 75 bars and 320 kelvins, "solution" clouds are present. At around 50 bars and 270 kelvins, water-ice clouds are present. Between 2 and 5 bars (100 and 150 kelvins respectively), clouds of solid particles of ammonium hydrosulfide form. Frozen ammonia crystal clouds form at around 5 bars and 120 kelvins. An ocean of ammonia dissolved in water has been postulated to exist below the 5-bar region. Frozen methane clouds occur at around 2 bars to 0.5 bar (70 kelvins to around 55 kelvins), according to predictive models. The temperature at the cloud tops is approximately 55.1 kelvins. The fraction of methane in the atmosphere at the 1-bar level is about 2 to 3 percent.



Neptune, with its high-altitude clouds, from Voyager 2. (NASA/JPL)



Voyager 2 images dark spots on Neptune; three are visible here, including one in the south with a white center. (NASA/JPL)

Absorption of infrared by methane contributes to the observed light blue color of Neptune. The fraction of methane drops significantly at higher altitudes but is still well above the ratio of methane to hydrogen found in the atmosphere of the Sun. Traces of carbon monoxide (CO) of about 1 part per million and hydrogen cyanide (HCN) of about 1 part per billion have been detected. The HCN is supposed to be formed through photochemical reactions with molecular nitrogen in the upper atmosphere. The abundance of CO in the stratosphere is greater than what can be explained by photochemical formation. One hypothesis is that CO comes from the lower regions. Another is that CO is formed when ice-bearing meteorites ablate in the upper atmosphere, with the methane participating in the reaction.

Neptune orbits the Sun once every 165 years but rotates on its axis once every 16.05 hours. The axis is inclined at 29°, so there are noticeable seasonal changes. Over the past few decades, the south pole has been closer to the Sun and is hence warmer. The cloud cover has become noticeably brighter, and there is evidence of strong updrafts.

Heidi Hammel and other researchers have concluded from analysis of narrow-angle images from Voyager 2 that large-scale cloud features in the equatorial and tropical regions move fast enough against the direction of rotation that their period of rotation about the axis is as high as 18.4 hours, compared to the planet's 16.05-hour day, which is deduced from radio signals based on the planet's internal rotation. This gives wind speeds of 325 meters per second, or 1,170 kilometers per hour. A large "Dark Spot" seen in the southern atmosphere by Voyager 2 appears to have been replaced by several dark spots in the south and later in the far northern hemisphere. These appear to be storms that form and dissipate. In the dark spots, wind speeds up to 2,400 kilometers per hour have been postulated. Several bright spots seen in infrared images appear to indicate warmer rising plumes of air that come from warm layers deep down but reach the upper atmosphere.

KNOWLEDGE GAINED

Much of what is known about Neptune comes from the Voyager 2 spacecraft, which transmitted images from an approach distance of 4.48 million kilometers in 1989. The Hubble Space Telescope in 1994-1995, mid-infrared data from the Keck telescope in Hawaii, the National Aeronautics and Space Administration (NASA) Mauna Kea telescope, and the Very Large Array radio telescope in New Mexico have added to our understanding of Neptune.

Predictive models use data from the observed emission and absorption through the atmosphere at various wavelengths of radiation. Sunlight reflected off clouds in the upper troposphere and lower stratosphere shows bright bands between 20° and 50° south. Distinct bright clouds are seen around 70° south. A bright south polar "dot" is visible. Observations in narrow bands corresponding to the absorption wavelengths of methane and ethane show that in the south polar regions, these substances appear to be less abundant. Scientists argue that this indicates a subsiding flow in the south polar region that draws the light gases down into warmer regions of the atmosphere and prevents them from freezing, as would happen if they rose into the upper atmosphere. The subsiding flow heats the lower levels adiabatically, while the lack of frozen crystals in the "dry" upper atmosphere above the south polar region makes the atmosphere transparent to much greater depths and allows the warmer regions below to be detected. It appears that air is rising more in the mid-southern and northern latitudes, and sinking at the equator and the south pole. This global circulation pattern is partially attributable to solar heating. Smaller clouds, such as the bands observed near 70° south latitude, are attributed to local weather.

A controversial theory for the intense atmospheric activity seen on this distant planet comes from Glenn Orton and his team, investigating an "electric universe theory." They have reported temperatures near Neptune's south pole that are high enough to allow gaseous methane from lower levels to escape into the upper atmosphere without freezing. They attribute this level of energy input to Neptune to an electrical connection with the rest of the solar system and out to the surrounding interstellar environment. Thus the heating is associated more with perturbations of the magnetic field of Neptune than with direct optical flow of thermal energy from the faraway Sun. According to this theory, Neptune is part of the "electric circuit" and hence can experience levels of energy input that are impossible to predict using just the distance from the Sun. Orton and his colleagues argue that conventional models of atmospheric dynamics based on solar heating fail when dealing with faraway planets.

CONTEXT

Expected to be much colder and less active than Uranus because it is 30 AU from the Sun, Neptune has surprised scientists with its profound seasonal changes in cloud cover and atmospheric circulation over the past twenty years. This is about half of a season. Hammel and colleagues conclude that cloud-top wind speeds are roughly the same order for all planets ranging from Venus to Neptune, even though solar energy inputs to their atmospheres differ by three orders of magnitude. Estimates of Neptune's internal energy and the amount of solar radiation that reaches it do not yet explain this level of activity. Radical theories of electric current flows from the Sun to the outer limits of the solar system have been advanced as alternative explanations. Neptune demonstrates, therefore, how the study of an extreme situation can challenge models based on Earth's atmospheric behavior.

Continued observation of Neptune with both ground-based telescopes and the Hubble Space Telescope must suffice until a Neptune orbiter similar to either the Galileo or Cassini designs might be dispatched to Neptune. In early 2002, such a probe, called the Neptune Orbiter Probe, was investigated but not funded—in part because of its projected cost and in part because it would have required nuclear propulsion to make a trip to Neptune in a reasonable amount of time. Such propulsion technology was not yet available. When the James E. Webb Space Telescope, an infrared observatory, enters service, it will join the instruments currently used to scrutinize the dynamics of Neptune's atmosphere.

Narayanan M. Komerath

FURTHER READING

Cruikshank, Dale P. *Neptune and Triton*. Phoenix: University of Arizona Press, 1995. Chapters are based on papers presented at an international conference of researchers and span all aspects of the Neptunian system based on Voyager images and data.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system, from early telescopic observations through the space missions that had investigated all planets (with the exception of Pluto) by the publication date. Takes an astrophysical approach to place the solar system in a wider context as just one member of similar systems throughout the universe.

Hammel, H. B. *The Ice Giant Systems of Uranus and Neptune*. New York: Springer, 2006. An authoritative discussion of the state of our knowledge of the Neptunian atmosphere and other aspects of both Uranus and Neptune, including the rings and moons of these planets.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science, this volume focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Accessible to serious readers.

Miner, Ellis D., and Randii R. Wessen. *Neptune: The Planet, Rings, and Satellites*. New York: Springer, 2002. Authors were members of the Voyager team during the Neptune encounter. This volume is another member of the Wiley-Praxis series in astronomy and astrophysics and is accessible to the general reader interested in planetary science.

See also: Auroras; Dwarf Planets; Earth's Atmosphere; Earth's Composition; Jovian Planets; Neptune's Great Dark Spots; Neptune's Interior; Neptune's Magnetic Field; Neptune's Ring System; Neptune's Satellites; Planetary Atmospheres; Planetary Classifications; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Pluto and Charon; Uranus's Atmosphere; Uranus's Interior.

Neptune's Great Dark Spots

Categories: The Neptunian System; Planets and Planetology

An elliptical atmospheric feature in Neptune's southern hemisphere—one large enough to contain Earth—was named the Great Dark Spot in 1989. It was interpreted as an anticyclonic storm. This dark spot eventually disappeared, but others subsequently appeared in the high northern latitudes. Given our understanding of terrestrial storms as driven by solar heating, and that Uranus, which is larger and closer to the Sun, did not show such activity until recently, such phenomena on distant Neptune were very surprising.

OVERVIEW

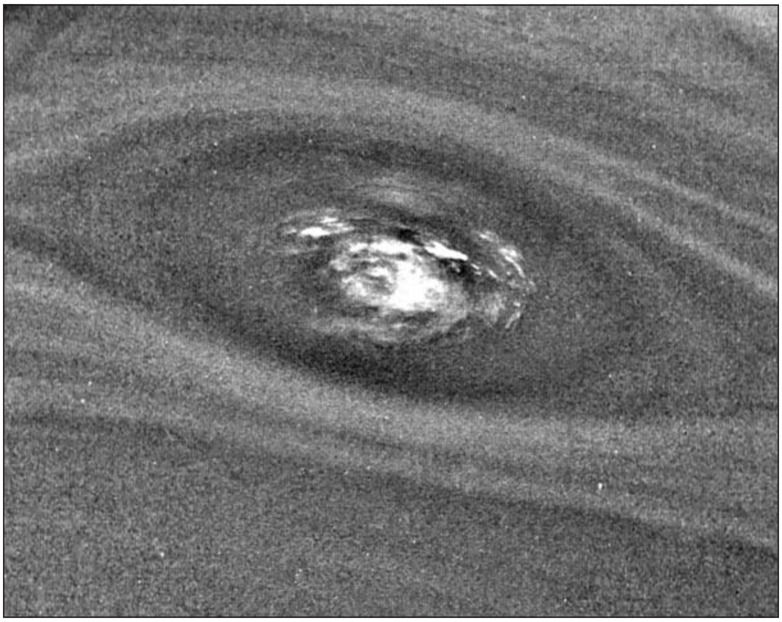
Images taken through filters that permit the passage of light at wavelengths of 467 nanometers were taken by the Voyager 2 spacecraft in 1989 at a distance of 2.8 million kilometers. Those images revealed an elliptic feature darker in albedo by about 10 percent. Covering between 30° and 45° of longitude and 8° to 17° of latitude, the Great Dark Spot (GDS) is about the size of Earth. It was initially positioned at 27° south latitude, drifting toward the equator at about 1.2° per month. Given the nomenclature GDS89, the Great Dark Spot was tracked from 27° to 17° south latitude over eight months and interpreted as a depression in Neptune's atmosphere, indicating a strong vortex such as a hurricane. GDS89 had a companion bright cloud band on one edge. This surprised astronomers after receiving featureless images of Uranus, a planet larger and closer to the Sun, during the Voyager 2 flyby in January, 1986.

In order to understand the concept of the Great Dark Spot, it is helpful to understand the vortex dynamics of storms. When gas heats up in the lower atmosphere and becomes less dense, it forms a "sinklike" flow, moving inward and then rising. The plume starts rotating, since conservation of angular momentum amplifies any difference in tangential speed as the radius decreases. In hurricanes that cover many degrees of latitude on a planet, the direction of rotation is predictable because of the Coriolis effect, whereby wind velocity interacts with the planet's rotation. The core remains clear of clouds and has rising warm air inside, while clouds revolve fastest around its periphery. As the plume reaches levels with lower density, the flow spreads out and the rotation slows outward. The warm, moist air (above Earth) condenses into clouds. Outside the core, cold air sinks. Thus, from observed cloud-top movement and temperature gradients, the strength of the storm and its axial upflow can be determined, along with the density differences and energy input that drive the storm. Prevailing regional winds and the Coriolis effect drive the storm across longitudes and toward the equator.

Near GDS89, wind speeds up to 2,400 kilometers per hour were recorded, but prevailing regional winds were on the order of 900 to 1,500 kilometers per hour. In addition to its drift, GDS89 showed oscillations of about 14° amplitude about the horizontal, with an eight-day period. Its aspect ratio fluctuated between 0.35 and 0.55, the longitudinal size varying between 30° and 45°, and the latitudinal extent between 12° and 17°. Over a 225-day period, GDS89 became increasingly circular. It also showed "tadpole-like tails," dark features off each of the smaller sides. A 1990 paper reported a dynamic model for such oscillations based on a single isolated vortex embedded in a shearing flow, and derived lower limits for the Rossby radius of deformation relating Coriolis forces and buoyancy forces.

In 1994, the Hubble Space Telescope (HST) observed that GDS89 had disappeared, but other dark spots appeared in the northern hemisphere. The Northern Great Dark Spot (NGDS32), a very stable dark spot, stayed near 32° north from around 1994 to 1996, and perhaps until 2000. Another dark spot, NGDS15, stayed near 15° north from March, 1996, to June, 1997. NGDS32 drifted across longitudes very steadily at about 36° per day. The drift rate of NGDS15, being much closer to the pole, proved harder to estimate. Bright methane clouds were associated with specific latitudes in any given period, but the active latitudes changed from –25° and –30° in 1989 to –30° and –46° in 1994-1996. Astronomers Heidi Hammel and G. Wesley Lockwood noted that GDS appeared in images where the dominant radiation was at 467 nanometers (blue), close to the brightest spots in the red (619 nanometers) and infrared (889 nanometers) existing on the planet at the time.

More detailed three-dimensional vortex simulations in 1998 provided explanations for the appearance of overlapping ellipses and correlated the drift rate of the spots with the prevailing wind speeds in the region. They predicted the breakup of these anticyclones near the equator, with Rossby waves propagating out over a few weeks. A 2001 paper in *Icarus* by P. W. Stratman et al. used simulations of storms, along with measurements of the bright accompanying clouds of the GDS, to estimate the atmospheric level where the top of the GDS



Voyager 2's close-up image of Dark Spot 2, with its white central structures, suggesting clouds. (NASA/JPL) should occur. If the top were in the stratosphere, the GDS would drift too quickly toward the equator and disperse. If the top were deep inside the troposphere, the clouds would be much larger than what was seen. Hence the top of a GDS should be near the tropopause. Based on this result, the pressure drop along a streamline threaded through a companion cloud was on the order of 3 millibars and the temperature change was on the order of one kelvin, indicating a lifting on the order of half a kilometer and relative wind speeds between the dark spot and the surrounding winds of 45 meters per second eastward.

However, a 2002 paper in *Icarus* by L. A. Sromovsky et al. showed that the steady latitudes maintained by NGDS32 and NGDS15 are not consistent with the model of anticyclonic storms, which should have moved strongly across latitudes. Computational fluid dynamic simulations by R. P. LeBeau and colleagues in 2006 and 2007 captured shape and oscillation phenomena similar to those of GDS89, but the size and oscillation amplitudes were off by as much as a factor of two. They calculated the shear profile in the background winds of Neptune that would be needed to explain the slow drift toward the equator of GDS89, but the much slower rates of the northern dark spots remain a challenge to explain.

In comparison, for more than three hundred years the Great Red Spot (GRS) of Jupiter has circled that planet along a southern latitude with minimal longitudinal drift. Attempts to explain the GRS as the flow around a high solid surface peak have been abandoned, and it is now considered to be a shallow cloud system trapped between shearing layers of horizontal winds. The redness relative to the surrounding white ammonia clouds indicates some temperature difference. Efforts to model the GRS as a vortical upwelling of fluid from below have met with limited success. It is not certain that the dark spots of Neptune are shallow structures, or that they extend to the cloud tops, rather than being features lying below clear atmospheric regions.

In 2006, the Hubble Space Telescope detected a dark spot more than 1,000 kilometers in extent, at 27° south latitude on Uranus. This observation was made as that planet began experiencing increased atmospheric activity with the coming of summer in its eighty-four-year orbit around the Sun. Scientists believe that Uranus is not as bland as Voyager 2's images suggested. Instead, as the amount of solar radiation that the planet intercepts increases, the planet may develop features seen in the atmospheres of the other gas giant planets.

KNOWLEDGE GAINED

What is known about the Great Dark Spots comes from images taken by the Voyager 2 spacecraft in 1989, by the Hubble Space Telescope from 1994 to 2000, and by ground-based optical and radio telescopes, including the Keck telescope, the Mauna Kea observatory, and the Very Large Array radio telescope in New Mexico. These are all passive observations, and the investigations for which they provide evidence depend on analyzing images taken with various filters that show specific wavelengths. Given the great distance, even the latest Hubble observations do not begin to approach the resolution achieved with Voyager 2's 1970's-vintage cameras. The narrow-angle camera on Voyager 2 was the instrument used to capture cloud-top images, which were then used to calculate wind speeds at that level.

Radio telescopes capture signals that should indicate rotation of the magnetic field and the internal structure of the planet, and hence give rotation rates based on those factors. With a planet composed mostly of fluid, there can be large differences between these rotation rates, which remain unexplained but are attributed to extremely high winds, which imply large frictional losses that require high energy input. The structure of the dark spots is derived mostly from simulations of fluid mechanics based on what is known of terrestrial storm systems and the limited data from these planets.

Researchers continue to model the dark spots using the fluid mechanics of hurricanes. Their speed of travel around the planet appears to match what is known or assumed of local wind speeds. However, their slowness in crossing latitudes requires fortuitous combinations of wind profiles to explain. Their shorter persistence compared to the Great Red Spot of Jupiter appears consistent with the much higher "wind speed" and the shear between different zones prevailing on Neptune. The spectral contrast suggests the presence of different gases and different temperature in the center of a spot, suggesting strong vertical motions of gas from the warmer depths of Neptune. The absence of features inside the dark spots, unlike the clouds seen above the GRS, frustrates efforts to derive their interior structure and vorticity directly.

CONTEXT

The driving engine for Neptune's dark spots remains puzzling, since the small changes in solar intensity associated with seasonal changes do not provide sufficient differences to explain these mysterious phenomena. If such strong weather activity can occur with such low intensity of sunlight, clearly much remains to be learned about why the Earth's weather behaves as it does. Observations of summertime on Uranus in the next decade may explain some of the mysteries, but Neptune's great dark spots are unlikely to yield their secrets until spacecraft descend through the Neptunian atmosphere to probe the planet's winds. New discoveries from such missions could greatly improve our ability to predict the course and evolution of killer storms on Earth.

In 2002 a study was convened to investigate the possibility of a Neptune Orbiter Probe, but the concept was not funded and was hampered by the need for nuclear propulsion technology, still to be developed. Until the development of a nuclear propulsion system capable of making the journey to Neptune in a timely fashion with a large payload, the likelihood of a follow-up spacecraft-based investigation of Neptune is relatively small. In the meantime, continuing study of Neptune and its atmospheric features will be conducted using the Hubble Space Telescope and ground-based observatories like the Keck telescope.

Narayanan M. Komerath

FURTHER READING

Cruikshank, Dale P. *Neptune and Triton*. Phoenix: University of Arizona Press, 1995. Chapters are based on papers presented at an international conference of researchers and span all aspects of the Neptunian system based on Voyager images and data.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system, from early telescopic observations through the space missions that had investigated all planets (with the exception of Pluto) by the publication date. Takes an astrophysical approach to give our solar system a wider context as just one member of similar systems throughout the universe.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college text on astronomy, somewhat more advanced than many introductory texts, with a wealth of detail and excellent diagrams. Chapters 6 through 16 describe the solar system. Comes with a CD-ROM.

Hammel, H. B., and G. W. Lockwood. "Atmospheric Structure of Neptune in 1994, 1995, and 1996: HST Imaging at Multiple Wavelengths." *Icarus* 129, no. 2 (October, 1997): 466-481. This paper uses the Hubble telescope to relate observations of dark spots on Neptune to the brightness of methane clouds. It shows that dark spots occur near the region where the brightest features appear in the red (619-nanometer) and near-infrared (889-nanometer) bands.

Hammel, H. B., G. W. Lockwood, J. R. Mills, and C. D. Barnet. "Hubble Space Telescope Imaging of Neptune's Cloud Structure in 1994." *Science* 268, no. 5218 (1995): 1740-1742. Discusses image results, wind speeds, the disappearance of the original dark spot, and the appearance of new ones. Also discusses methods used to filter the images to bring out different features based on the absorption of red and infrared by methane, and the bright scattering from high-altitude clouds above the gaseous methane layers.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Accessible to serious readers.

Miner, Ellis D., and Randii R. Wessen. *Neptune: The Planet, Rings, and Satellites*. New York: Springer, 2002. The authors were members of the Voyager team during the Neptune encounter. Accessible to the general reader interested in planetary science.

Oxlade, Chris. *Jupiter, Neptune, and Other Outer Planets*. New York: Rosen Central, 2007. Intended for middle school students, this book compares and contrasts the Jovian planets using the latest data available on the gas giants and their features.

See also: Auroras; Dwarf Planets; Earth's Atmosphere; Earth's Composition; Jovian Planets; Neptune's Atmosphere; Neptune's Interior; Neptune's Magnetic Field; Neptune's Ring System; Neptune's Satellites; Planetary Atmospheres; Planetary Classifications; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Pluto and Charon; Uranus's Atmosphere; Uranus's Interior.

Neptune's Interior

Categories: The Neptunian System; Planets and Planetology

Neptune is a gas giant planet, so its interior structure is completely different from that of the terrestrial planets and most similar to that of Uranus, another gas giant. The most likely structure consists of a relatively small rocky core, an icy layer, and a molecular hydrogen layer. The interior is then covered with an atmospheric layer of hydrogen, helium, and various other gases. Neptune's interior lacks the metallic hydrogen layer found in Jupiter and Saturn.

OVERVIEW

It is not possible to conduct direct studies of planetary interiors, including Earth's, so astronomers must resort to indirect clues to infer the interior structures and compositions of planets. These clues can include characteristics of the bulk density, the planetary magnetic field, seismic activity and waves for terrestrial planets, and heat flow from the interior.

The two major classes of planets in the solar system are the Jovian planets and the terrestrial planets. Jupiter is the prototype for the Jovian planets, which also include Saturn, Uranus, and Neptune. Earth is the prototype for terrestrial planets, which also include Mercury, Venus, and Mars. Earth's moon is also very similar in structure to the terrestrial planets. In oversimplified terms, Jovian planets are big balls of gas, and terrestrial planets are small balls of rock and metal.

A planet's bulk density is one clue to its interior composition and structure. Density is defined as the mass divided by the volume. Both an object's mass and its volume depend on the object's composition and how much of that material there is. For example, a boulder has a greater volume and mass than a pebble made of the same type of rock. However, the amount of material cancels out when dividing the mass by the volume. Therefore the density of a boulder and the density of a pebble made of the same type of rock will be the same, even if the boulder is as large as a planet. Density is a property of the type of material but not how much material there is, so the density of a planet can tell us something about its composition. Water has a density of 1.000 kilograms per meter cubed (km/m^3). Planets with approximately this density are primarily icy materials or gas. Planets with a density of approximately 2,000 to 3,000 km/m^3 are typically made of rocky materials. Planets with a density of 7,000 to 9,000 km/m^3 would be primarily metallic in composition. Planets with densities between these ranges are mixtures of materials. Neptune has a density of 1,600 km/m^3. It must therefore be mostly ice or gas. This density, the highest of the Jovian planets, is slightly greater than the density of Uranus and significantly greater than the densities of Jupiter and Saturn. Therefore, the rocky cores of Neptune and Uranus constitute a larger portion of their total mass than do the cores of Jupiter or Saturn. Because Neptune's density is slightly greater than that of Uranus, the two planets must have differences in their interior compositions, despite their very close similarities.

Seismic activity can give us clues to the interiors of only the terrestrial planets (which are solid and therefore have seismic activity). Neptune is a gas giant planet with no potential for seismic activity, so this potential clue does not apply to Neptune.

Because the Jovian planets are gas, astronomers can use the ideal gas law for theoretical calculations of interior properties. This equation relates the temperature, density, and pressure of the gas.

Planetwide magnetic fields give us clues to the planets' interiors. All magnetic fields are ultimately produced by some type of electric current, so in order for a planet to have a magnetic field, its interior must contain a liquid or gas layer that can conduct electricity. The planet's rotation helps set up the electric currents in this layer. A composition of iron or other ferromagnetic materials can enhance the magnetic field produced by the interior electric currents but is not essential. Neptune does have a strong magnetic field that is comparable in strength to those of Saturn and Uranus. Like that of Uranus, Neptune's magnetic field axis is tilted significantly from its rotational axis. The tilt is 46° in the case of Neptune. The center of Neptune's magnetic dipole is displaced significantly from the center of the planet. This displacement is even greater than for Uranus. No other planets have such displacements or tilts of their magnetic dipoles. If the icy layer is not completely solid, but at least partially molten, then the water can conduct electric currents and produce a magnetic field. Impurities in the water increase its ability to conduct electricity. Hence we infer that Neptune's icy layer is not completely solid and the molten regions are not symmetric about the center of the planet.

With the exception of Uranus, the Jovian planets emit more energy than they receive from the Sun. Neptune emits about three times as much energy as it receives from the Sun. This extra energy must come from somewhere, and the only place available is the planet's interior. Therefore the interiors of the Jovian planets, including Neptune, must be hot. When a gas giant planet forms, its own gravity compresses the gas. A gas heats up when it is compressed, so the interiors of gas giant planets should be hot. Neptune therefore has a hot interior, as expected from the behavior of gases and confirmed by the heat flowing outward from the interior to the surface. The existence of large storms and weather on Neptune provides additional evidence that heat energy, which is needed to power storms, flows upward from the interior.

Putting all these clues together, astronomers can infer the interior structure of Neptune. The innermost layer is a rocky core. Neptune's rocky core is approximately the size of Earth and has a temperature of about 7,000 kelvins. The mass of Neptune's rocky core is most likely a little less than ten times the mass of Earth, but estimates range from four to fifteen times Earth's mass. The rocky core is more massive than Earth, even though it is about the same size, because the high temperatures and pressures in the core of the planet increase the density of the rock.

Above the rocky core, Neptune has an icy mantle consisting primarily of water, methane, and ammonia ice that is not completely frozen. The ice is slushy rather than completely frozen. Hence convection currents can flow in this mantle. The convection currents carry heat from the interior to the surface. They also help form Neptune's magnetic field. As the slushy material circulates, it is moving in the Sun's extended magnetic field. This motion induces electric currents in the slush. The slushy water ice has ammonia dissolved in it, so it conducts electricity very well. These induced electric currents induce Neptune's magnetic field. Electric currents induce magnetic fields and moving or changing magnetic fields induce electric currents. If the regions where the magnetic field is generated are far from the center of the planet and distributed asymmetrically around the planet, then the magnetic field would be tilted and displaced as the Voyager 2 data show.

The outermost layer of Neptune's interior is molecular hydrogen. The temperature, pressure, and density of the molecular hydrogen layer increases with depth. The icy layer can form despite high temperatures because the freezing temperature of water increases as the pressure increases.

Neptune's interior structure is similar to that of Uranus but dissimilar from the interiors of the other gas giants, Jupiter and Saturn, which have a metallic hydrogen layer in addition to these layers. Neptune and Uranus do not have metallic hydrogen layers because the pressure in the hydrogen layers is not high enough for hydrogen to become metallic hydrogen rather than molecular hydrogen.

KNOWLEDGE GAINED

Most of what is known about Neptune and its interior comes from the Voyager 2 mission. After flying by Jupiter and Saturn, Voyager 1 flew out of the plane of the solar system. Voyager 2 flew on to Uranus and Neptune. The Neptune flyby was in 1989. To date, this is the only spacecraft to explore either Uranus or Neptune. All other observations are from ground-based or orbiting telescopes. Examples of the knowledge of Neptune gained from the Voyager mission include measuring the planet's magnetic field, taking more accurate measurements of Neptune's mass and therefore its density, and discovering its surface storms.

The Great Dark Spot that the Voyager discovered on Neptune is a large storm, similar in nature to the great red spot on Jupiter. Since the Voyager mission, the Hubble Space Telescope has been able to take pictures with enough resolution to allow astronomers to follow weather patterns on Neptune. Storm systems on Neptune change. Weather and large storm systems require energy to drive them. On Earth the energy comes from the Sun, but Neptune is too far from the Sun for solar energy to drive the storms and weather astronomers observe. Therefore the needed energy must come from the interior of the planet. Uranus, though very similar to Neptune, does not have such storms. It is deduced, therefore, that Neptune has a much hotter interior than Uranus.

The interior differences among the four Jovian planets provide with clues to their formation. Jupiter, Saturn, Uranus, and Neptune all have rocky cores of approximately the same size and mass. Jupiter and Saturn, however, have much more hydrogen and helium surrounding their rocky cores. This observation suggests that Uranus and Neptune formed much later than Jupiter and Saturn, at a time when the protoplanetary disk had dissipated more of its hydrogen and helium gas. At the greater distance from the Sun, the small planetesimals that had to merge to form the cores of Uranus and Neptune were farther apart and therefore took longer to merge.

Neptune, like the other planets, is differentiated into layers. The denser materials are closer to the center. Differentiation in solid terrestrial planets tells us that they were at one time liquid, because denser materials would sink only in a fluid, not in a solid. Smaller satellites that were never liquid are not differentiated.

CONTEXT

Comparing the interiors of other planets casts light on Earth's interior, just as knowledge of Earth's interior provides insights into other planets. These comparisons of how planetary interiors develop provide clues to the origins of the planets and to the processes that shape planetary interiors and compositions.

To date, only one spacecraft, Voyager 2, has studied the outer planets Uranus and Neptune. No probes have yet flown by Pluto. This paucity of robotic probes to the outer solar system limits our understanding of these planets. When humans send additional spacecraft to the outermost reaches of the solar system, our understanding will increase. When we send another probe to Neptune, our knowledge of the planet and its interior will increase dramatically. It is very likely to completely change our ideas about Neptune's interior.

Paul A. Heckert

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. One chapter of this readable introductory astronomy textbook covers Neptune, Uranus, and Pluto.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. Chapter 14 of this introductory astronomy textbook is a complete overview of the outermost planets in the solar system.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. This textbook on the satellites and planets of the solar system summarizes our understanding of planetary interiors in chapter 8.

Hester, Jeff, et al. *Twenty-First Century Astronomy*. New York: W. W. Norton, 2007. Chapter 8 of this well-illustrated astronomy textbook is about the Jovian planets and includes a good discussion of their planetary interiors.

Morrison, David, Sidney Wolf, and Andrew Fraknoi. *Abell's Exploration of the Universe*. 7th ed. Philadelphia: Saunders College Publishing, 1995. The outer planets are covered in chapter 16 of this classic astronomy textbook.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. Cambridge, England: Cambridge University Press, 2002. An extremely well written introductory astronomy textbook. Chapter 10 is an overview of our knowledge of the Jovian planets.

Zeilik, Michael, and Stephen A. Gregory. *Introductory Astronomy and Astrophysics*. 4th ed. Fort Worth, Tex.: Saunders College Publishing, 1998. Aimed at undergraduate physics or astronomy majors, this textbook goes into more mathematical depth than most introductory astronomy textbooks. Chapter 6 covers the basic principles of the Jovian planets.

See also: Auroras; Dwarf Planets; Earth's Atmosphere; Earth's Composition; Jovian Planets; Neptune's Atmosphere; Neptune's Great Dark Spots; Neptune's Magnetic Field; Neptune's Ring System; Neptune's Satellites; Planetary Atmospheres; Planetary Classifications; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Pluto and Charon; Uranus's Atmosphere; Uranus's Interior.

Neptune's Magnetic Field

Categories: The Neptunian System; Planets and Planetology

Neptune is the eighth planet outward from the Sun. Although classified a Jovian or "gas giant" planet, Neptune shares more similarities with Uranus than it does with Jupiter or Saturn. Neptune's magnetic field is inferred to be generated in a manner akin to that which produces Uranus's magnetic field.

OVERVIEW

The first planet to be discovered during the telescope age was Uranus, the seventh planet from the Sun. The planets from Mercury (closest to the Sun) out through Saturn (sixth from the Sun) were known to the ancients. No records exist that pinpoint a discovery event in each of these planets' cases. After Sir William Herschel announced the discovery of Uranus in 1781, observers began making precise determinations of Uranus's orbit about the Sun.

Unexpected motions in the orbit of Uranus led many to suspect that a planet existed farther out in the solar system, one that exerted gravitational influences on Uranus that perturbed its orbit. A search followed for a new planet to join the family of seven planets in the contemporary solar system model. This search was greatly guided by mathematical analysis using the young field of celestial mechanics, as based on an understanding of Newtonian gravitation. The basis of that analysis involved calculating where a body would have to be located beyond Uranus in order to account for measured variations in the orbit of Uranus.

Who actually discovered Neptune has been a matter of debate. Calculations in 1843 by John Couch Adams were dispatched to Astronomer Royal Sir George Airy, but the latter failed to be convinced. He even asked Adams to send proof of the validity of his work. Adams appeared to be either disappointed or insulted by Airy's dismissal, and he never responded. Two years later, Urbain Le Verrier published an independent calculation that did generate some interest in Airy. Efforts at Cambridge Observatory failed to find the eighth planet, but when Le Verrier asked the Berlin Observatory to aim a telescope toward a particular region of the sky, Johann Gottfried Galle reacted immediately on the request. Neptune was found on September 23, 1846. Its position was just one degree off from Le Verrier's calculation and twelve from Adams's. In the aftermath, the British and French argued over who actually discovered the new planet. Adding to the controversy was the fact that Cambridge records indicated Neptune was charted months earlier than Galle's observation. However, those records did not recognize Neptune as a planet. Even in the last years of the twentieth century, there were those who argued against giving Le Verrier and Adams joint credit as Neptune's discoverers.

In the thirteen decades since its discovery, information about Neptune has been garnered only by using ever-increasingly large Earth-based telescopes. Then, in 1977, a special opportunity arose for sending a single spacecraft to explore the outer solar system making use of gravity assists in turn from Jupiter to Saturn to Uranus to Neptune. The National Aeronautics and Space Administration (NASA) was able to launch Voyager 2 on such a "Grand Tour." Voyager 2 encountered Jupiter in 1979, Saturn in 1981, Uranus in 1986, and then Neptune in 1989.

Just as it had at Uranus, Voyager 2 would answer the very basic question of whether or not Neptune had a magnetic field. The answer, be it in the affirmative or the negative, would provide considerable insight into the interior structure of Neptune. Voyager 2 indeed found Neptune to be a dynamic world that generates more heat energy than the radiation it receives from the Sun. It generates a complex and dynamic magnetic field that has left scientists with much to consider concerning the implications of that magnetic field for the structure and evolution of the planet.

KNOWLEDGE GAINED

There are a number of means whereby a magnetic field can be detected. Most of them make use of electromagnetic induction in that a voltage is induced in a coil by intercepting a time-dependent magnetic flux within that coil. That is the basis of sophisticated magnetometers such as that flown on the Voyager 2 spacecraft. Scientists had to be careful in interpreting readings from the spacecraft's magnetometer, in that it picked up magnetic effects created by the solar wind even at the distant position of Neptune. Solar wind particles become trapped within a planet's magnetosphere, and that generates its magnetic field.

The Voyager 2 Neptune encounter began on June 5, 1989, with the spacecraft still 117 million kilometers distant from the planet. Voyager first detected radio waves, indicating it had crossed into Neptune's magnetosphere, on August 24. The spacecraft came within 4,400 kilometers of Neptune's upper atmosphere on August 25. The spacecraft remained inside Neptune's magnetosphere for a total of thirty-eight hours and noted that Neptune's magnetic "bubble" was affected by both the planet's ring system and its satellites. The planet's magnetosphere exists for 35 Neptune radii on the sunward side and out to 72 radii behind the planet. That varies with the strength of the solar wind. Several days later the Neptune near-encounter phase ended after Voyager passed close to Neptune's large satellite Triton. Triton especially alters the outermost portion of Neptune's magnetosphere. Triton revolves about Neptune in a retrograde fashion. Charged particles within the magnetosphere near Triton also include nitrogen ions that originate from cryogeyser eruptions on this satellite.

Voyager 2 also found radiation belts trapped in the planet's expansive magnetosphere, although charged particle density in Neptune's magnetosphere was found to be less than that at Uranus. The composition of the trapped particles primarily includes protons, electrons, and ionized molecular hydrogen and helium.

Analysis of Voyager 2 data from both the spacecraft's magnetometer and radio astronomy experiment indicated that Neptune's magnetic field varies considerably as the planet rotates in the presence of the solar wind. Neptune's magnetic field is dominated by its dipole character, but it also has a strong quadrupole component. A quadrupole can be understood as if it were two bar magnets oriented at right angles to each other. The field strength near the equatorial region is 1.42 microteslas. Neptune's dipole strength is 2.2×10^{17} tesla-meters cubed. Uranus's dipole moment is nearly double that value. The planet's quadrupole moment is due in large part to an offset of the field center from the planet's center. Octupole and higher moments could not be determined accurately.

The center of Neptune's field is displaced from the planet's center by approximately 0.55 planetary radius, which corresponds to 55 percent of the planet's radius. The planet's magnetic field is also tilted significantly, 47° relative to the planet's rotational axis, as was the case at Uranus.

On Earth, trapped particles spiral around magnetic field lines and dip down into the atmosphere over polar regions, creating auroras by exciting atmospheric gases. Voyager 2 detected auroral activity in Neptune's atmosphere, but in large part because of the complexity of Neptune's magnetic field structure, the observed auroral activity was not confined to areas near the magnetic poles of the planet.

Observations of Neptune's magnetic field provided insight into Neptune's interior. Like Uranus, Neptune is believed to have a liquid layer of high electrical conductivity in motion powered by internal heat flow from the core beneath. This results in a dynamo effect from this fluid mantle region, which is composed of water, ammonia, methane, and lesser amounts of volatile substances. Planetary rotation results in complex variations of Neptune's magnetic field.

Neptune is a radio source as a result of its rotation and the fact that it possesses a reasonably strong magnetic field. Rotation of the magnetic field is determined by the observed periodicity in the planet's radio emissions. That provided a better means of determining Neptune's rotation rate than attempting to monitor the time taken for atmospheric features, such as streamers or large storms, to make a compete rotation. A gas giant's atmosphere rotates differentially. The planetary core is involved in producing the magnetic field, and radio emissions led scientists to determine Neptune's core rotates once every 16.11667 hours.

CONTEXT

After Uranus, Neptune was the second planet in the solar system to be discovered by telescopic observations. Just as the search for Neptune originated based on irregularities noticed in the orbit of Uranus, perturbations of Neptune's orbit led to a search for yet another planet in the outer solar system. In 1930, Pluto was discovered by Clyde Tombaugh, but ironically Pluto did not explain Neptune's orbital irregularities. Also, as of 2006, the International Astronomical Union no longer classifies Pluto as a planet, but rather Pluto is now considered to be either a dwarf planet or the first representative of a class of objects called plutoids.

For more than one hundred years since the planet's discovery, the only means to investigate Neptune was by using ever-larger Earth-based optical telescopes. Then, with the advent of the space age and the confluence of a special configuration of planets in the outer solar system that arises only once every 176 years, it became possible to send a spacecraft from Earth: Voyager 2. It traveled in turn to Jupiter, Saturn, Uranus, and Neptune. The Voyager 2 flyby of Neptune generated the best available data and images about this mysterious blue world to that time.

After that Voyager flyby, the Hubble Space Telescope was used in the 1990's and 2000's to make continual observations of Neptune. However, at this time only a new spacecraft mission can advance our knowledge of Neptune's magnetic field. A proposal advanced in 2002 for the Neptune Orbiter Probe was investigated in 2003 as part of NASA's response to the Bush administration's Vision for Space Exploration but was not funded at that time. Planetary scientists continued to indicate strong support for a mission, in the class of the Cassini orbiter, to investigate both Uranus and Neptune. The need to develop nuclear propulsion to get to the outer solar system, however, will need to be met first.

David G. Fisher

FURTHER READING

Bredeson, Carmen. *NASA Planetary Spacecraft: Galileo, Magellan, Pathfinder, and Voyager*. New York: Enslow, 2000. This book is part of Enslow's Countdown to Space Series. Provides an overview of NASA planetary exploration during the last two decades of the twentieth century. Designed for younger readers but suitable for all nonspecialist audiences.

Cruikshank, Dale P. *Neptune and Triton*. Phoenix: University of Arizona Press, 1995. Chapters are based on papers presented at an international conference of researchers and address all aspects of the Neptune system based on Voyager images and data.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system, from early telescopic observations through the space missions that had investigated all planets (with the exception of Pluto) by the publication date. Takes an astrophysical approach to place our solar system in a broader context as just one member of similar systems throughout the universe.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college text on astronomy, somewhat more advanced than many introductory texts, containing a wealth of detail and excellent diagrams. Chapters 6 through 16 describe the solar system. Comes with a CD-ROM.

Hunt, Garry E., and Patrick Moore. *Atlas of Neptune*. Cambridge, England: Cambridge University Press, 1994. A combination of telescope and Voyager images of the planet is supplemented by text on historical and scientific background. Summarizes the contemporary understanding of Neptune. Useful to general readers and scientists alike.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Accessible to the serious general audience.

Kerrod, Robin. *Uranus, Neptune, and Pluto*. New York: Lerner Publications, 2000. Presents investigations of the outer solar system from Pioneer 11 through Voyager 2. Intended for a young audience. Not all photographs are clearly identified, but the text is suitable for the general audience.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar-system astronomy suitable for an introductory college course as well as nonscientists. Filled with supplemental learning aids and solved student exercises. A Web site is available for educator support.

Miner, Ellis D., and Randii R. Wessen. *Neptune: The Planet, Rings, and Satellites*. New York: Springer, 2002. The authors, who were members of the Voyager team during the Neptune encounter, have written an book accessible to the general reader interested in planetary science.

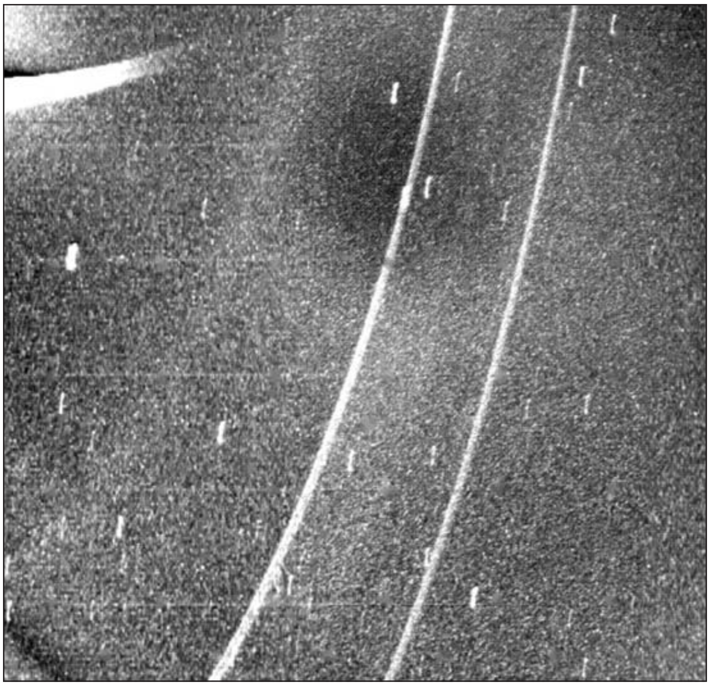
Oxlade, Chris. *Jupiter, Neptune, and Other Outer Planets*. New York: Rosen Central, 2007. Intended for middle school students, this work compares and contrasts the Jovian planets.

See also: Auroras; Dwarf Planets; Earth's Atmosphere; Earth's Composition; Jovian Planets; Neptune's Atmosphere; Neptune's Great Dark Spots; Neptune's Interior; Neptune's Ring System; Neptune's Satellites; Planetary Atmospheres; Planetary Classifications; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Pluto and Charon; Uranus's Atmosphere; Uranus's Interior.

Neptune's Ring System

Categories: The Neptunian System; Planets and Planetology

While all of the gas giant planets have ring systems, Neptune's are unique. They consist of a series of concentric rings and partial arcs. Why Neptune's rings developed differently is still a matter of debate among scientists.



After reaching Neptune in 1989, Voyager 2 took this image of rings around Neptune. (NASA/JPL)

OVERVIEW

After Uranus's discovery, scientists noticed that it did not behave as expected. They concluded that gravitational effects of a yet unknown planet caused discrepancies noted in Uranus's orbit. In the early 1840's, British astronomer John Adams and French mathematician Urbain Le Verrier independently calculated the unknown planet's orbit and mass. In 1846, German astronomer Johann Galle discovered Neptune within one or two degrees of its predicted location.

It was not until 1984, however, that scientists at two observatories found the first evidence of a ring system around Neptune. On July 22 of that year, Neptune occulted star SAO 186001. Astronomers use the occultation of stars to look for rings and satellites or to measure a planet's exact size. The observatories (which were located in Chile about 100 kilometers apart) recorded a brief occultation lasting just a second. The starlight was reduced by a mere 35 percent when Neptune passed across the line of sight from Earth to the star. Usually this would be considered evidence of a new satellite. In this case, however, to fit with the data, a satellite would only have been only 10 to 20 kilometers in diameter. The astronomers concluded instead that they had discovered a ring around Neptune. However, neither observatory noticed any reduction in starlight on the other side of Neptune.

Neptune's arc-type ring structures remained a mystery until the Voyager 2 spacecraft arrived at Neptune in 1989. The first images from Voyager were taken at a distance of about 20.8 million kilometers during Voyager's approach to Neptune. The photographs showed two arcs around the planet. The first arc was about 48,000 kilometers long and tilted at a 45° angle with respect to Neptune. The inner arc was about 9,600 kilometers wide, at a distance of approximately 51,680,300 kilometers from Neptune's center. Both arcs were found near recently discovered satellites of the planet. Voyager 2 located the partial rings early enough for scientists at the spacecraft's headquarters in California to redirect its camera toward them. Initial photographs showed a differing brightness in the rings, which scientists thought might have resulted from size, density, or rock variations among the ring particles. Saturn's F ring has a similar varying brightness.

On August 23, 1989, Voyager 2 transmitted images back to Earth that solved the partial ring arc mystery. The arcs are actually full rings that encircle Neptune but are not completely visible from Earth. The spacecraft photographed the rings in both forward and backward scattered light. Microscopic particles cannot be seen from Earth, since they are too small to reflect enough sunlight to be seen at such a distance. However, when they were backlit—imaged with the Sun behind them—Voyager 2's cameras were able to detect them as small dust particulates.

In total, Neptune has five rings and partial arcs. The innermost ring, named Galle, is inside the orbit of Neptune's satellite Naiad. This ring is composed mostly of dust, resembling the par-

Rings of Neptune

	Radius (km)	Radius/ Eq. Radius	Optical Depth	Albedo (×10^{-3})	Width (km)
Neptune equator	24,766	1.000	—	—	—
Galle (1989N3R)	~41,900	1.692	~0.00008	~15	~2,000
LeVerrier (1989N2R)	~53,200	2.148	~0.002	~15	~110
Lassell (1989N4R*)	~53,200	2.148	~0.00015	~15	~4,000
Arago (1989N4R*)	~57,200	2.310	—	—	<~100
Unnamed (indistinct)	61,950	2.501	—	—	—
Adams (1989N1R)	62,933	2.541	~0.0045	~15	~50
Arcs in Adams Ring:					
Courage	62,933	2.541	0.12	—	~15
Liberté	62,933	2.541	0.12	—	~15
Egalité 1	62,933	2.541	0.12	~40	~15
Egalité 2	62,933	2.541	0.12	~40	~15
Fraternité	62,933	2.541	0.12	—	~15

Notes: *Lassell and Arago were originally identified as one ring, designated 1989N4R.
Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.

tial arcs. The Le Verrier ring is next closest to Neptune. It is also the second most prominent of the rings. It is narrow and dusty, located 700 kilometers outside the orbit of Despina. The widest ring, Lassell, extends for 4,000 kilometers. It is one of the less dusty and more complete rings. Its outer edge, which is significantly brighter than the rest, has been named the Arago ring. The outer and most prominent ring is named after one of Neptune's discoverers, Adams. Orbiting Neptune about 1,000 kilometers outside the path of Galatea, it is narrow and dim in comparison to the bright rings of Uranus and Saturn. The Adams ring also contains five ring arcs: Courage, Liberté, Égalité (1 and 2), and Fraternité.

After Voyager 2 reached Neptune, astronomers knew that the planet had a complete ring system. They still did not know, however, what had caused the arcs of the Adams ring to remain incomplete. In 1991, planetary scientist Carolyn Porco postulated that the arcs were a result of Neptune's signature satellite Galatea. The small satellite orbits Neptune 1,000 kilometers inside the Adams ring. Porco argues that Galatea acts as a shepherding moon, whose gravitational effects keep the matter in the arcs from forming a complete ring. Analysis of data collected by Voyager 2 shows that the arcs "wiggle," or shift positions, by up to 30 kilometers. Porco believes arc distortion occurs at the right speed to be attributed to the small satellite. Critics of this theory point out that ring-arc material would need to have orbits that intersect each other in order to maintain its overall shape. Porco agrees that this would cause collisions leading to the inevitable destruction of the arcs themselves. Her model also only partly explains the possible origin of the arcs. It can only determine places where arcs are more likely to develop, not specifically where they already have developed. Because of the lumpiness of the partial rings, Porco and other scientists believe they could once have been a small moon that was destroyed.

In 1999, a group of scientists argued that Galatea cannot be the sole influence causing Neptune's arcs. They studied data collected in 1998 using the Hubble Space Telescope. The largest change in position was of the Liberté arc, which was displaced 1.9° compared to the Égalité arc. The scientists argued that this finding was not a result of varying particle size, given Voyager data that support size conformity within the arc. Their findings, however, do not rule out the shepherding moon theory that relies on the effects of two satellites. Porco and her colleague Fathi Namouni published a paper in 2002 again arguing in favor of their Galatea shepherding-moon model. They believe that Galatea's elliptical orbit keeps the arc particles from having many collisions. By refining the mathematical model of Neptune's ring system, Porco and Namouni proved that Galatea might still be the answer to the arc mystery. More precise data concerning Galatea's orbit and the partial rings are needed before scientists will be able to determine what really causes the arcs around Neptune.

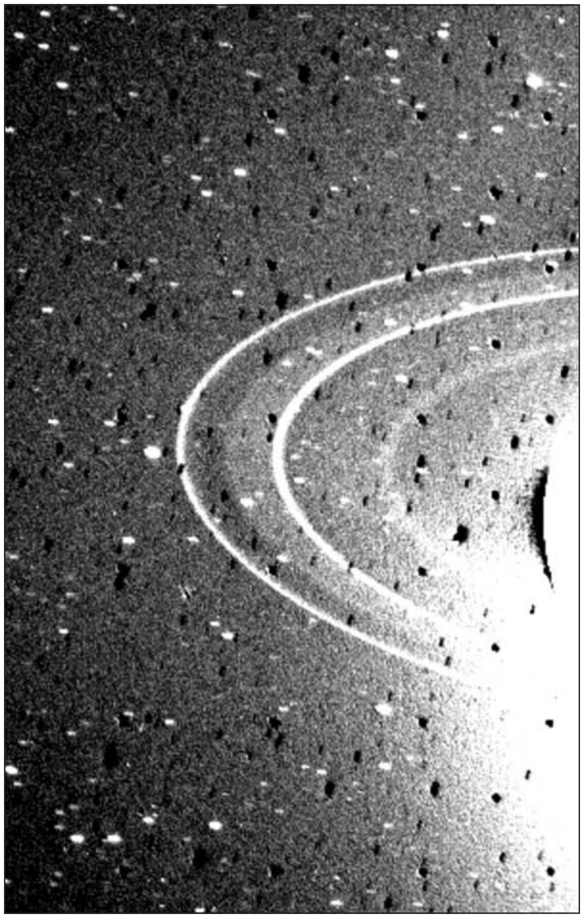
In 2002 and 2003, a group of scientists led by Imke de Pater photographed Neptune's outer rings. Studying the images, they discovered that all of Neptune's arcs appear to be fading away. The Liberté arc showed the most deterioration when compared with the Voyager data. If the dissipation continues at its current rate, the Liberté arc will be gone within one hundred years. Their observation proved that whatever is holding the ring arcs together is not regenerating them fast enough to sustain them.

KNOWLEDGE GAINED

Several stellar occultations with Neptune were observed, but only five showed evidence of a ring or arc system. Scientists working at the European Southern Observatory and the Chilean Cerro Tololo Observatory both witnessed the same brief occultation of star SAO 186001. The two teams noticed that the planet itself did not block the star, because the starlight diminished by only about 35 percent. Neither group was able to locate any reduction of light on the opposite side of Neptune. These data led to the conclusion that Neptune had only a partial ring or arc system.

The Voyager program consisted of two spacecraft focused on studying the outer planets of our solar system. Launched in 1977, Voyager 2 reached Neptune twelve years later. The spacecraft had two video cameras, infrared and ultraviolet spectrometers, and other instruments. It was not until Voyager 2 started its approach toward Neptune that the first photographs of the rings and arcs were taken. Scientists learned that Neptune in fact did have a mostly complete ring system, like the rest of the gas giants, but that the dust particles that composed the majority of the rings were too small to be detected by Earth-based telescopes. Full rings could be seen only when backlit by the Sun—a view scientists can get only with the use of space probes such as Voyager.

Scientists have more recently been able to use the Hubble Space Telescope's Near Infrared Camera and Multi-Object Spectrometer (NICMOS) to study Neptune's ring system. The camera was powerful enough to detect ring arcs during two different occultations in 1998. One of the difficulties in studying them from Earth is the visual proximity of the rings to Neptune itself. NICMOS solved this problem by using a



A wide-angle image from Voyager 2, captured in 1989, shows two main Neptunian rings and a faint, inner ring, behind which stars shine. The long exposure resulted in the brightness of Neptune's disk. (NASA) special filter that blocks wavelengths at which methane, a main component of Neptune's atmosphere, reflects light.

Starting around the year 2000, astronomers were able to use Earth-based telescopes to study Neptune's ring system. Imke de Pater and his colleagues viewed Neptune's ring arcs using the 10-meter Keck Telescope in Hawaii. Advances in image resolution and light-gathering power made this possible. However, the Keck can still detect only the brighter Adams ring; the others remain too faint for it to observe. When the Keck data were analyzed, scientists found evidence that the ring arcs are fading away, with the Liberté arc showing the most damage. If the arcs continue to degrade at the rate they have since 1989, within a hundred years they will have disappeared.

CONTEXT

All of the gas giants in our solar system have ring systems. They all share similarities and have their differences as well. Neptune's system contains five full rings and five arcs. Until telescope technology advances enough to detect the faint dust particles that compose the rings, only the brightest, outermost ring, the Adams ring, will be able to be photographed from Earth.

Learning more about the ring systems promises to help scientists better understand the origins and evolution of the solar system itself. Scientists are also still debating what is preventing Neptune's arcs from forming complete rings. The two main competing theories are based on the idea of shepherding moons that gravitationally trap the particles in arcs. Further advancements must be made before more can be learned about Neptune's ring system, and sending another spacecraft, with more advanced instruments, will be the best way to do so. However, Voyager reached Neptune only through the "gravity assists" of the other planets, which were aligned in a way that occurs only once every 176 years. Short of waiting another century for the planets to realign, the only way to send a new probe to Neptune (preferably one in the Cassini class) would be to develop nuclear propulsion systems that are still in the research and development phases.

Jennifer L. Campbell

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. This well-written college-level text, designed for introductory astronomy courses, includes a chapter on Uranus and Neptune that covers the ring systems.

Esposito, Larry. *Planetary Rings*. Cambridge, England: Cambridge University Press, 2006. A synopsis of current knowledge of the outer planets' ring systems. Includes information from the Cassini mission and on Neptune's rings and arcs, as well as ring ages and evolution. Geared toward scientists and college students.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. An introductory college text that gives students easy-to-understand analogies to help explain complex theories. Includes a CD-ROM featuring InfoTrac software.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A thorough and well-written college-level introductory astronomy book. Covers all aspects of Neptune, including its ring systems.

Fridman, Alexei M., and Nikolai N. Gorkavyi. *Physics of Planetary Rings: Celestial Mechanics of Continuous Media*. New York: Springer, 1999. Compares the ring systems of Jupiter, Saturn, Uranus, and Neptune using observational and mathematical data. Ideal for scientists, students, or amateur astronomers wishing to know more about the rings of the outer planets.

Miner, Ellis D., and Randii R. Wessen. *Neptune: The Planet, Rings, and Satellites*. New York: Springer, 2002. Covers Voyager 2, its mission, difficulties, and discoveries. Includes a chapter dedicated to Neptune's ring system. Written for nonscientists.

Miner, Ellis D., Randii R. Wessen, and Jeffrey N. Cuzzi. *Planetary Ring Systems*. New York: Springer Praxis, 2006. Looks at the ring systems of each gas giant. Covers recent research in the field, as well as the many questions that remain unanswered.

Tabak, John. *A Look at Neptune*. London:

Franklin Watts, 2003. Discusses Voyager 2's mission to Neptune, focusing on what the spacecraft discovered about the planet and what remains unknown. A good introductory work for general audiences.

See also: Auroras; Dwarf Planets; Earth's Atmosphere; Earth's Composition; Jovian Planets; Neptune's Atmosphere; Neptune's Great Dark Spots; Neptune's Interior; Neptune's Magnetic Field; Neptune's Satellites; Planetary Atmospheres; Planetary Classifications; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Pluto and Charon; Uranus's Atmosphere; Uranus's Interior.

Neptune's Satellites

Categories: Natural Planetary Satellites; The Neptunian System; Planets and Planetology

Neptune's family of thirteen known satellites has challenged astronomers to push the limits of their observational technology and theoretical knowledge. While similar in some respects to the satellite systems of the other three Jovian planets, Neptune's satellites, especially Triton, have proven to have unique characteristics, which allow astronomers to increase their understanding of conditions in the early solar system.

OVERVIEW

Eight days after Neptune's discovery on September 23, 1846, astronomer John Herschel sent a letter to colleague William Lassell, suggesting that he might turn his telescope toward the new planet to search for satellites. Nine days later, on October 10, Lassell discovered what would become known as Triton, the largest satellite of Neptune and the seventh largest satellite in the solar system, with a diameter of 2,706 kilometers. The satellite's name is credited to French astronomer Camille Flammarion and was used unofficially for decades before being officially adopted. Mythologically, Triton was the son of the Greek sea god Poseidon (Roman name, Neptune).

Triton proved to be puzzling to astronomers; unlike most large satellites, it orbits its planet retrograde, or "backward" (opposite to the direction of Neptune's rotation). Its orbit is nearly circular but is highly inclined to Neptune's equator (23°). These anomalous features led astronomers to hypothesize that Triton was captured by Neptune's gravity during a close approach to the planet, despite the fact that other planetary satellites suggested to be likewise captured, such as Saturn's Phoebe, are much smaller.

Despite several searches, no other satellites were discovered around Neptune until the Dutch-born astronomer Gerard Kuiper found a faint (magnitude 19.5) satellite using the 82-inch telescope at McDonald Observatory in 1949. The name Nereid was suggested, after the fifty sea-nymph daughters of Nereus and Doris, attendants of Poseidon in Greek mythology. As strange as Triton's orbit was, Nereid's was no less peculiar. Although it orbits Neptune in a prograde direction at an average distance some fifteen times greater than that of Triton, its orbit is highly elongated (eccentric), with its most distant orbital point nearly seven times farther than its closest approach to the planet. Such an orbit is more similar to that of a comet than the orbits of most satellites, so—although at a diameter of 340 kilometers Nereid is larger than most asteroids—some astronomers suggested that it also was captured by Neptune's gravity. Others pointed an accusatory finger at Triton, suggesting that the capture of Triton might have disrupted Nereid's original orbit. The discovery of Nereid thus led to more questions than answers, some of which awaited the flyby of the Voyager 2 spacecraft in late August, 1989.

Nereid was too distant (4.7 million kilometers) to be effectively imaged by Voyager 2, which has proven to be a lingering source of frustration to astronomers. The brightness of the satellite is known to change on both short and long timescales, which astronomers have interpreted as due to the rotation of the satellite (although no definitive period can be established). The changes in brightness may be due to differences in the surface material on different sides of the satellite (as in the case of Saturn's satellite Iapetus) or a nonspherical shape. Interestingly, water ice has been spectroscopically found on the surface, leading to speculations that the satellite's surface might be a mixture of ice and some dark material, similar to the satellites of the other Jovian planets (and dissimilar to known Kuiper Belt objects). In addition, Halimede, the innermost of the outer five irregular satellites, has spectroscopic similarities to Nereid, and their orbits have a 0.41 in 1 probability of collision during the history of the solar system. This has led to the suggestion that Halimede is a splinter of a nonspherical Nereid.

The other three Jovian planets were found to have many satellites, which could be divided into two broad classes: regular moons, usually larger in size with prograde, nearly circular orbits nearer the planet; and more distant, irregular satellites, usually smaller and with greater orbital eccentricities and inclinations. Irregular satellites appear to be captured objects, while regular satellites were considered to be the "original" satellites of the planet, dating from the formation of the planet itself. Since Voyager 2 had discovered three new satellites around Jupiter, four at Saturn, and ten orbiting Uranus, it was not surprising when the Voyager 2 imaging team announced the discovery of a new Neptunian satellite in June, 1989, two months before the probe's closest approach to Neptune. Now named Proteus for a Greek sea god, it is actually larger than Nereid (416 kilometers in diameter), although it is exceedingly difficult to observe from earthbound telescopes because of its close orbit (roughly a third Triton's distance from Neptune).

Three more satellites, all roughly one-third the size of Proteus and orbiting closer to Neptune, now named Larissa, Galatea, and Despina, were discovered in late July, 1989. It was later determined that Larissa had been the cause of a brief dip in brightness seen in a star that Neptune had nearly occulted in May, 1981, an observation that had initially been explained as due to the planet's then undiscovered rings. Thalassa and Naiad, each a fraction of the size of Proteus, were discovered orbiting even closer to Neptune only a few days before Voyager's August 25, 1989, closest approach. The names were selected from mythological characters associated with Poseidon/Neptune, according to the International Astronomical Union's (IAU's) convention for Neptunian satellites and based on the historical precedence of Triton and Nereid. For example, Despina was a daughter of Poseidon and Demeter, and Galatea was one of the Nereids. All six of the new satellites were "regular" in the sense that they had prograde orbits that were nearly circular and, with the exception of Naiad, were well aligned with Neptune's equator.

Following the Voyager 2 flyby, numerous attempts were made to discover additional satellites orbiting farther from Neptune than Nereid, which led to the discovery of five "irregular" satellites in 2002 and 2003: in order out from Nereid, these were named Halimede, Sao, Laomedeia, Psamathe, and Neso, all named after individual Nereids in Greek mythology. Sao and Laomedeia have prograde orbits, the other three satellites orbit retrograde, and all five satellites have significant orbital eccentricities (0.3-0.6). Each is between approximately 40 and 60 kilometers wide, comparable to Naiad or smaller, and are largely presumed to have been captured by Neptune.

KNOWLEDGE GAINED

Even though Neptune was found to have both "regular" and "irregular" satellites, significant mysteries remained after their discovery. The five innermost satellites were found to be potato-shaped (not unexpected for small satellites whose self-gravity is small). However, some astronomers felt this was more evidence of violence in the Neptunian system. The large size of Triton and its surprising orbit suggested that it had played an important role in the past and future dynamics of the entire satellite system. In 1966, Thomas McCord calculated these effects and found that Triton is in an unstable orbit that eventually will lead to its destruction in possibly tens of millions of years (or more). When it passes inside Neptune's Roche limit, tidal effects will tear it to shreds. McCord also suggested that Nereid's peculiar orbit was caused by the effects of Triton's capture and later orbital evolution from an initially parabolic or nearly parabolic orbit to a nearly circular one.

Later researchers continued to use various computer simulations and computational techniques to model the evolution of the Neptunian system, in an attempt to explain both the gross and subtle characteristics of each moon's orbit. Many astronomers presume that Neptune had an original generation of satellites, possibly similar to the satellites of Uranus, which was destroyed in the aftermath of Triton's capture, either through mutual collisions between the satellites or through satellites being gravitationally slingshot out of the Neptunian system. The debris is posited to have formed a disk, some of which was accreted by Triton, increasing its diameter and mass, while some of the material accreted to form the currently observed six regular (inner) satellites. The irregular (outer) satellites are suggested to be survivors from this catastrophic event. An unknown number of other satellites collided with each other or one of the larger satellites, or were slingshot out of the system. In the process, the orbits of some of the satellites that were originally in the retrograde-type orbits (normally seen in irregular satellites) were changed into prograde orbits. The orbits of all thirteen satellites have slowly evolved since then through mutual gravitational effects, and these orbits continue to evolve today.

The detailed images taken of Triton by Voyager 2 strengthened the assumption that it was originally a Kuiper Belt object (similar to Pluto). One of the other important discoveries of the Voyager 2 Neptune flyby was Neptune's five rings, named (in order out from Neptune) Galle, Le Verrier, Lassell, Arago, and Adams. As in the case of the other Jovian planets, astronomers expected the ring system and the inner moons to have gravitational interactions. It is suggested that Galatea affects the particles of the Adams ring, which orbit just beyond the orbit of Galatea, although the exact nature of the gravitational resonance is still being debated. Interestingly, it is the Adams ring that has the famous "arcs" (which led to its nickname, the "sausage ring"). Despina, Thalassa, and Naiad orbit between the Le Verrier and Galle rings. Despina has been suggested to be acting as a shepherding moon.

Astronomers continue to explore the details of the satellites' orbital dynamics through cutting-edge computer simulations and theoretical models, and in so doing test our assumptions about the evolution of the solar system in general and the Jovian planets in particular.

CONTEXT

The Voyager 2 mission led to an explosion of knowledge about the Neptunian system, as it did for the Jovian, Saturnian, and Uranian systems before it. However, given the late discovery of the regular satellites, Voyager 2 was able to target only the outermost four of the eight then-known satellites: Triton, Nereid, Proteus, and Larissa.

Little is known about the composition of the other satellites, although they are presumed to have low densities (0.4-0.8) similar to that of Saturn's small satellites Janus, Epimetheus, and Prometheus. Since the exact orbital interactions between the satellites depends on their densities, further research into those dynamics will allow astronomers to put further constraints on the satellites' densities, and hence compositions, and any future observational evidence gathered on the satellites' compositions will lead to further refinements in models of the Neptunian system's dynamic history. Therefore, although Voyager 2 has certainly been a boon to Neptune researchers, further significant breakthroughs certainly await any future spacecraft visiting Neptune and its thirteen satellites.

Kristine Larsen

FURTHER READING

Cruikshank, Dale P., ed. *Neptune and Triton*. Tucson: University of Arizona Press, 1995. This thick tome is a collection of the invited papers presented at a 1992 scientific conference summarizing what was known about Neptune and its satellites following two years of analysis of the Voyager 2 flyby data.

Cuk, Matija, and Brett J. Gladman. "Constraints on the Orbital Evolution of Triton." *Astrophysical Journal* 626 (2005): 1113-1116. This technical paper details the results of a computer model of the effects of Triton's capture on preexisting satellites.

Holman, Matthew J., et al. "Discovery of Five Irregular Moons of Neptune." *Nature* 430 (2004): 865-867. A technical summary of the methods used to discover the five outermost satellites of Neptune with earthbound telescopes.

McCord, Thomas B. "Dynamical Evolution of the Neptunian System." *Astronomical Journal* 71, no. 7 (1966): 585-590. The first detailed study of the history and future dynamics of Triton's orbit.

Miner, Ellis D., and Randii R. Wessen. *Neptune: The Planet, Rings, and Satellites*. Chichester: Springer, 2002. This book summarizes most of the important results and papers on the Neptunian system through its publication date. The Voyager 2 mission and spacecraft are described in detail.

Moore, Patrick. *The Planet Neptune: An Historical Survey Before Voyager*. 2d ed. New York: John Wiley and Sons, 1996. A popular historical account of the discovery of Neptune, Triton, and Nereid from a famous amateur astronomer and science writer. The epilogue extends the book through the major discoveries of the Voyager 2 mission.

Zhang, Ke, and Douglas Hamilton. "Orbital Resonances in the Inner Neptunian System. I. The 2:1 Proteus-Larissa Mean-Motion Resonance." *Icarus* 188 (2007): 386-399. A technical review of computer simulations of past gravitational interactions between Proteus, Larissa, Triton, and Neptune, with the goal of explaining the current orbital parameters and estimate the densities of Proteus and Larissa.

_______. "Orbital Resonances in the Inner Neptunian System. II. Resonant History of Proteus, Larissa, Galatea, and Despina." *Icarus* 193 (2008): 267-282. A technical investigation of the orbital histories of these satellites as well as estimates of the satellites' densities (and therefore compositions).

See also: Dwarf Planets; Eris and Dysnomia; Jupiter's Satellites; Miranda; Neptune's Atmosphere; Neptune's Great Dark Spots; Neptune's Interior; Neptune's Magnetic Field; Neptune's Ring System; Planetary Satellites; Pluto and Charon; Triton; Uranus's Satellites.

Neutrino Astronomy

Category: Scientific Methods

Neutrinos—nearly massless particles with no charge—are generated in nuclear reactions, such as those deep in the interior of the Sun and other stars or in stellar explosions. Neutrinos do not readily interact with matter, which makes their detection difficult, but their detection is important, because it provides a way to probe basic energy-generation processes in stars and stellar explosions.

OVERVIEW

Neutrinos are uncharged, massless or nearly massless, particles, first proposed in 1930 by physicist Wolfgang Pauli to account for the apparent missing energy in beta decay. Beta decay is the emission of an electron from the nucleus of some radioactive elements. To balance energy but manage to go unobserved in the experiments that had been conducted, Pauli's hypothetical neutrino had to interact only very weakly with ordinary matter. For more than thirty-five years, particle physicists were unable to confirm the existence of Pauli's elusive particle experimentally.

Efforts to detect the neutrino rest on observing one of the consequences of "inverse beta decay," a process in which a neutrino collides with a proton and they are transformed into a neutron plus a positron. Neutrino detectors can function by using a photomultiplier to detect the Cherenkov light emitted by the relativistic positron, or by using scintillation counters to detect gamma rays given off by annihilation of the positron when it encounters an electron. Crude estimates of the likelihood of inverse beta decay reactions suggest that a single neutrino would have to pass through a solid wall many billions of kilometers thick to have, on the average, just one such interaction. These estimates indicate that neutrino detectors have to be very large in order to detect a single neutrino interaction event.

In 1956, physicists Clyde L. Cowan and Frederick Reines, working at the Los Alamos National Laboratory, built the world's largest scintillation counter, a device to detect the small flash of light given off by the interaction of a neutrino in the apparatus. When placed adjacent to the Savannah River Nuclear Reactor, a high-power reactor that produces about 10^{18} neutrinos per second, their apparatus detected only one neutrino interaction every twenty minutes. Nevertheless, Cowan and Reines were able to confirm the existence of Pauli's elusive particle.

Thermonuclear reactions occurring in the core of the Sun also produce neutrinos. High temperatures and pressures in the Sun's core permit nuclear reactions in which two hydrogen nuclei combine to form a deuterium nucleus, a positron, and a neutrino. This process, called the "p-p reaction," also releases a considerable amount of energy. The resulting deuterium nuclei can then react to form helium and again release energy causing the core temperature to rise. This process allows the helium nuclei to combine to form an even heavier nucleus. Two of these reactions are important to the production of neutrinos inside the Sun. In the "beryllium-electron capture" reaction, a beryllium 7 nucleus combines with an electron to form lithium 7 and a neutrino. Eventually, this lithium 7 will react with a proton to produce boron 8. The boron 8 can then decay into two helium nuclei, a positron, and a neutrino.

Eventually, fusion reactions in a star will terminate when the core has been converted into iron, since reactions that produce even heavier elements from iron all use energy rather than releasing it. At this stage, if the mass of the star is large enough, the core collapses under its own weight, converting protons into neutrons and emitting a burst of neutrinos. The result is a supernova explosion, accompanied by an intense but short burst of neutrino emission and the formation of a neutron star. Efforts to model the supernova explosion process indicate that, although a bright flash of visible light can be observed for months after the explosion, about 99 percent of the energy is carried away by the neutrinos.

Neutrinos produced in the core of the Sun, because of their low probability of interaction, can pass entirely through the Sun and reach Earth unimpeded. Their number and energies carry information about temperature, pressure, and nuclear reaction processes in the core.

There are three major neutrino-producing reactions in the Sun: the p-p reaction, which produces neutrinos with a continuous range of energies up to about 420,000 electron volts (eV); the beryllium-electron capture reaction, which produces neutrinos with a single energy of 862,000 eV (this reaction produces about thirteen times fewer neutrinos than the p-p reaction); and the boron-decay reaction, which produces neutrinos with a continuous energy distribution up to about 15 million eV (this reaction produces about one ten-thousandth as many neutrinos as the p-p reaction). Although the boron-decay reaction produces far fewer neutrinos than either of the other two reactions, the high energy of these decay neutrinos makes them much easier to detect on Earth.

Efforts to develop an apparatus for neutrino astronomy have concentrated on detecting the continuous flux of neutrinos emitted by nuclear reactions in the Sun and short bursts from supernova events. Theoretical models of both processes have predicted the number and energies of the neutrinos to be expected from both sources.

At the same time that Cowan and Reines were attempting to detect the neutrino, a chemist at Brookhaven National Laboratory, Raymond Davis, Jr., began experiments that eventually led to the detection of neutrinos from the Sun. Rather than employing an electronic detector, which counted individual events as they occurred, Davis used an idea, originally suggested in 1948 by Bruno Pontecorvo, for a chemical detector that could accumulate the chemical by-product of neutrino reactions over several months. Pontecorvo's idea rested on the fact that a neutrino interacting with a neutron in an atom of chlorine 37 transforms that nucleus into a nucleus of argon 37. It was hoped that argon, being a chemically inert gas, could be extracted easily from large volumes of chlorine-rich liquid. Since argon 37 is radioactive, it could be identified easily by its decay following extraction. Between 1954 and 1956, Davis tested a version of this detector, containing 1,000 gallons of chlorine-rich carbon tetrachloride, located near a nuclear reactor at the Brookhaven

National Laboratory. In the scientific paper describing their results, Davis suggested that this detector could be scaled up in size to permit detection of solar neutrinos. The chlorine 37 reaction, however, requires energetic neutrinos, so it is sensitive only to the high-energy neutrinos produced by the boron-decay reaction in the Sun.

After several pilot projects to demonstrate the effectiveness of the radiochemical techniques for neutrino detection, Davis constructed a solar neutrino observatory buried about 2 kilometers below the surface in the Homestake Gold Mine in Lead, South Dakota. The neutrino detector consisted of a large cylindrical tank, 6 meters in diameter and 15 meters long, containing 100,000 gallons of perchloroethylene (containing 520 tons of chlorine). The tank was surrounded by water to shield the detector from neutrons emitted by trace quantities of uranium and thorium in the walls of the mine. The detector was first placed into operation in May, 1967, and was allowed to sit passively accumulating argon atoms from neutrino interactions for periods of one to three months. Then, Davis would extract the argon produced during the accumulation period and determine the argon 37 abundance by counting its decays. Using this apparatus, Davis succeeded in observing neutrinos from the boron-decay reaction in the Sun. His detector, however, provided neither directional information nor time resolution.

To sort out the signals from a supernova would require a directional instrument with real-time response. For these efforts, astrophysicists returned to electronic detectors traditionally employed in particle physics but greatly enlarged because of the low probability for neutrino interactions. These detectors consist of large tanks of water serving as targets for neutrino interactions by inverse beta decay. The water tanks are surrounded by photomultiplier detectors that see the tiny flash of Cherenkov radiation emitted by each relativisitic positron as it traverses the water. A consortium of physicists from the University of California at Irvine, the University of Michigan, and Brookhaven National Laboratory (IMB) operates a 7,000-ton water tank detector in a mine just outside Cleveland, Ohio. A second such detector, the Kamiokande II detector, containing 3,000 tons of water, is operated by Japanese physicists in the Kamioka Mine about 300 kilometers west of Tokyo. Both of these detectors recorded neutrino emissions from Supernova 1987A, providing information on the energy distribution and flux of neutrinos for comparison with theoretical models of the explosion process.

Neither the chlorine detector operated by Davis nor the IMB or Kamiokande II water and Cherenkov detector is sensitive to the lowest-energy neutrinos, which are those from the p-p reaction. To detect these low-energy neutrinos, astrophysicists are employing the chemical detection technique. The target is gallium, since gallium 71 can interact with a neutrino to produce germanium 71 and an electron. The germanium can be separated from the gallium, providing a measure of the neutrino flux. A neutrino detector consisting of 20 tons of pure gallium was commissioned at Baksan in the Soviet Union in early 1990. Like Davis's chlorine detector, this detector, as well as a similar one planned for operation in Gran Sasso, Italy, provides no information on the time or direction of the neutrino. Nevertheless, its sensitivity to low-energy neutrinos will permit detection of the neutrinos produced by the p-p reaction in the Sun.

APPLICATIONS

Astrophysical models of the Sun prominent in the 1950's and 1960's led Davis to suspect that the neutrino detector would record many events each month. To describe his results, Davis had defined a new unit, the solar neutrino unit (SNU), with one SNU corresponding to the production of a single atom of argon 37 in his apparatus every six days. Theoretical calculations in 1963 by astrophysicist John N. Bahcall, using the then current astrophysical models of the reactions occurring in the core of the Sun as well as the accepted physical properties of the neutrino, gave an expected capture rate of 50 SNUs.

Initial results from Davis's detector, from a 40-day run in May and June, 1967, and a 110-day run in June through October, 1967, were startling. Davis placed an upper limit of 3 SNUs, almost one-twentieth the expected amount, on the solar neutrino flux, but his results were consistent with detecting no solar neutrinos at all. Initially, astrophysicists suspected that the experiment was flawed, but Davis undertook a series of calibration experiments to establish the efficiency of his detector. He concluded that the experiment was not flawed. Bahcall repeated the theoretical calculations in 1969, using revised values for the nuclear reactions in the Sun. The new results lowered the expected flux to 8 SNUs; however, Bahcall could not explain the much lower results reported by Davis.

Davis improved his argon decay detector in 1970, increasing the sensitivity of the experiment. This improvement allowed the first detection of solar neutrinos. Although it lacked directional sensitivity in the detector, the source could only be inferred, because the Sun was expected to be the brightest continuous neutrino source in the sky. Davis's measurements continued almost uninterrupted from 1967 to 1990, giving essentially the same results over the entire period. The long-term average measured for more than twenty years is only 2.3 solar neutrino units, well below the most recent theoretical predictions. The low solar neutrino detection rate observed by Davis was confirmed by the Japanese Kamiokande II neutrino detector, which began operation in January, 1986. Since the Kamiokande II detector is an electronic detector, providing real-time information on the neutrino's direction, the Japanese physicists were able to confirm, as had been suspected, that most of the neutrinos detected were from the Sun.

Two general ways have been proposed to explain the large discrepancy between the flux of boron decay neutrinos from the Sun and the number expected from the astrophysical models. Either the physicists' basic understanding of the neutrino and its interactions is wrong or the astrophysical models of the interior of the Sun are incorrect. Physicists have suggested that the neutrino might have a small mass, which would alter some of the calculated reaction rates, or that the neutrinos emitted by the Sun might change into a second type of neutrino, not able to participate in the inverse beta-decay reactions on which the detectors rely. Astrophysicists, on the other hand, have suggested that, since the boron-decay reaction is a minor reaction in the core of the Sun, the real test of the solar model will come with the results from the gallium detectors, which will be sensitive to the p-p reaction.

A resolution to the solar neutrino problem came from the Sudbury Neutrino Observatory (SNO) in Canada in 2001. That detector incorporated 1,000 metric tons of heavy water (water where the hydrogen is replaced by deuterium) and 10,000 photomultiplier tubes located in a nickel mine 2 kilometers underground. SNO results were verified by the Super Kamiokonde neutrino detector array. A ramification of neutrinos having even a small value of mass is that they can change flavor. Electron neutrinos produced in nuclear reactions deep in the Sun's core could transform into muon and tau neutrinos before reaching Earth. Previous neutrino detectors were only capable of registering neutrinos of the electron flavor. Although SNO could detect the other flavors, it could not precisely determine the neutrino mass. The mass was estimated to be between 0.05 and 0.18 eV. Apart from solving the solar neutrino problem, another ramification of neutrinos having mass is that they may account for some, but not all, of the missing mass in cosmological models.

Neutrinos from supernovae were expected to be of much higher energy than those from the Sun and clustered very tightly in time. The first extrasolar neutrino event to be observed was Supernova 1987A, an explosion in the Large Magellanic Cloud. Both the IMB and Kamiokande detectors saw neutrino bursts, with eight neutrinos detected by IMB in a seven-second window and eleven detected by the Kamiokande II detector in a thirteen-second window. Based on the detector efficiency, 3×10^{16} neutrinos from Supernova 1987A passed through the IMB detector, resulting in eight interactions.

The number and energies of the neutrinos detected from Supernova 1987A were in general agreement with theoretical predictions. The Kamiokande II investigators indicate that the temperature of the explosion, as inferred from the neutrino energy distribution, is consistent with supernova models. Also, the time that elapsed between the neutrino burst and the optical brightening was consistent with model estimates for the blue giant progenitor star. The Japanese investigators calculated that neutrinos carried off 8×10^{52} ergs of energy, much more than the 10^{49} estimated to have been carried by the electromagnetic radiation. This finding confirms theoretical models showing that most of the gravitational energy liberated in the stellar collapse is carried away by neutrinos. If these data had been analyzed in real time, they would have provided an early warning for the optical astronomers, since the neutrino burst was observed three hours before the first optical sighting of the explosion.

CONTEXT

Neutrino astronomy, although still in its infancy, has called into question basic ideas about how the Sun generates its energy and has confirmed models of the supernova explosion process.

The discrepancy between the minimum number of solar neutrinos expected theoretically and the actual number detected by Davis was referred to as the "solar neutrino problem." Explanations for this discrepancy between theory and experiment have focused on two areas: Either the astrophysical models of the Sun are wrong or the physicists' understanding of the neutrino and its interactions is wrong. Modifications to the models for the Sun would require that the temperature or the helium content of the core is lower than modeled. On the other hand, the neutrinos emitted by the Sun may change type (from electron neutrinos to muon neutrinos) on their way from the Sun to the Earth. Since the detectors are sensitive only to the electron neutrinos, they would miss the muon neutrinos. The Sudbury Neutron Observatory and the Kamiokande neutrino detectors produced data that indicated that this solar neutrino problem was resolved as a result of neutrino oscillation. Earlier detectors could only pick up the electron flavor, but electron-flavor neutrinos transform into the muon and tau flavors. When that oscillation was taken into account, the validity of solar models was verified; the solar neutrino problem was no longer an issue, as the observed neutrino flux matched theoretical models.

Observation of a neutrino burst from Supernova 1987A was a more successful application of the theory and was consistent with theories of stellar explosions. Measurements demonstrated that most of the energy liberated in the gravitational collapse of a supernova is carried off by neutrinos. The study of these particles is important to the understanding of the supernova process.

The success of the early experiments in neutrino astronomy spurred proposals for other, more advanced instruments. The Japanese group proposed a Super Kamiokande detector, containing 32,000 tons of water. This detector would be able to register so many solar neutrinos from the boron-decay reaction that it could monitor the temperature of the solar core with an accuracy of 1 percent over weekly intervals. Also, the Super Kamiokande detector would be able to detect about four thousand neutrinos from each supernova occurring within the Milky Way, including events near the galactic center, which are shielded from optical observation because of the large amount of material along that direction. Such observation would provide information on the dynamics of star collapse. Super Kamiokande began operation in 1996. Then in 2001, more than 6,600 of its photomultiplier tubes imploded in a chain reaction. Efforts to reengineer the design to preclude such an unfortunate accident were set in motion. By 2006, Super Kamiokande II was back in operation.

Other detectors, including the Ga-Ge detector operated by Russia, will focus their efforts on observing the low-energy neutrinos from the Sun. Since the neutrinos are able to pass through the Sun unimpeded, they provide the only way to observe directly the nuclear processes occurring in its core.

George J. Flynn

FURTHER READING

Bahcall, John N. "The Solar-Neutrino Problem." *Scientific American* 262 (May, 1990): 54-61. Explains how the nuclear processes in the Sun produce neutrinos and how they are detected. Summarizes the results of two decades of measurements by Davis. Discusses the implications of the missing solar neutrinos and describes the various solutions, such as neutrino oscillations, variations in the composition of the interior of the Sun, or a lull in solar activity. Well illustrated.

Fukugita, Masataka, and Tsutomu Yanagida. *Physics of Neutrinos*. New York: Springer, 2003. A good starting point for learning about neutrinos. Material is technical, heavily referenced, accessible primarily to scientists, especially particle physicists.

Giunti, Carlo, and Chung W. Kim. *Fundamentals of Neutrino Physics and Astrophysics*. New York: Oxford University Press, 2007. Both theoretical and experimental aspects of neutrinos are presented. Implications for astrophysics explained. Exhaustive references. Technical.

Helfand, David. "Bang: The Supernova of 1987." *Physics Today* 40 (August, 1987): 24-32. Nonmathematical discussion of the observation of Supernova 1987A. Describes the theory that neutrinos should be associated with supernovae and covers the detection of the neutrinos from the 1987 event.

Koshiba, Masa-Toshi. "Observational Neutrino Astrophysics." *Physics Today* 40 (December, 1987): 38-42. Well-illustrated, nontechnical article, describing the results from the Japanese Kamiokande II neutrino detector, including measurements of the solar flux and the burst of neutrinos from Supernova 1987A. Describes neutrino experiments planned for the 1990's around the world.

Serway, Raymond A., et al. *College Physics*. 7th ed. New York: Brooks/Cole, 2005. A textbook used at the introductory level in college physics courses. Filled with sample problems. Comes with PhysicsNow, an online teacher/student resource.

Thornton, Stephen T., and Andrew Rex. *Modern Physics for Students and Engineers*. 3d ed. New York: Brooks/Cole, 2005. A comprehensive presentation of the development of relativity, quantum mechanics, and theoretical and experimental aspects of nuclear and particle physics. For undergraduates or serious scientific researchers.

Tipler, Paul A., and Ralph Llewellyn. *Modern Physics*. 5th ed. New York: W. H. Freeman, 2007. A classic text in modern physics. Covers astrophysics, nuclear physics, and elementary particle physics at various levels of mathematical rigor. Homework problems could be better, but online aids exist.

Wolfenstein, Lincoln, and Eugene W. Beier. "Neutrino Oscillations and Solar Neutrinos." *Physics Today* 42 (July, 1989): 28-36. Describes attempts to measure the number of solar neutrinos arriving at Earth and examines the implications of the low neutrino flux measured for accepted models of the Sun and the neutrino.

Woosley, S. E., and M. M. Phillips. "Supernova 1987A!" *Science* 240 (May 6, 1988): 750-759. Describes the major observations of Supernova 1987A and explains how the neutrino observations are related to optical observations and supernova models.

Young, Hugh D., and Roger A. Freedman. *University Physics with Modern Physics*. 11th ed. New York: Addison Wesley, 2003. An undergraduate text that spans classical mechanics, thermodynamics, Maxwell's electrodynamics, optics, and modern physics. Includes a good first introduction to aspects of elementary particle physics as well.

See also: Coordinate Systems; Earth System Science; Gravity Measurement; Hertzsprung-Russell Diagram; Infrared Astronomy; Optical Astronomy; Radio Astronomy; Telescopes: Ground-Based; Telescopes: Space-Based; Ultraviolet Astronomy; X-Ray and Gamma-Ray Astronomy.

Novae, Bursters, and X-Ray Sources

Category: The Stellar Context

Many stars undergo violent changes in activity during their evolution as their chemical and physical makeups change. These stars—which include novae, supernovae, neutron stars, and black holes—release tremendous quantities of electromagnetic radiation (including X rays) during their violent phases.

OVERVIEW

The universe consists of perhaps 1 trillion galaxies. Each of these galaxies contains on the average approximately 100 billion stars, assuming that the universe is homogeneous. A star is a massive accumulation of hydrogen and helium atoms initially formed from the big bang explosion that created the universe, between 13 and 14 billion years ago. An average star such as the Sun, which is classified as a G2 star, contains the mass of more than 1 million Earths. Inside a star, the intense pressures and temperatures generated by the gravitational attraction of the hydrogen and helium atoms results in the fusion, or combining, of these atoms. Effectively, a star is 100 million hydrogen (fusion) bombs exploding every second. Hydrogen fuses to form helium during the star's early life. Once hydrogen fuel is exhausted, helium is fused to form beryllium, lithium, and heavier elements. This process of thermonuclear nucleosynthesis continues as the progressively older star burns heavier and heavier elements.

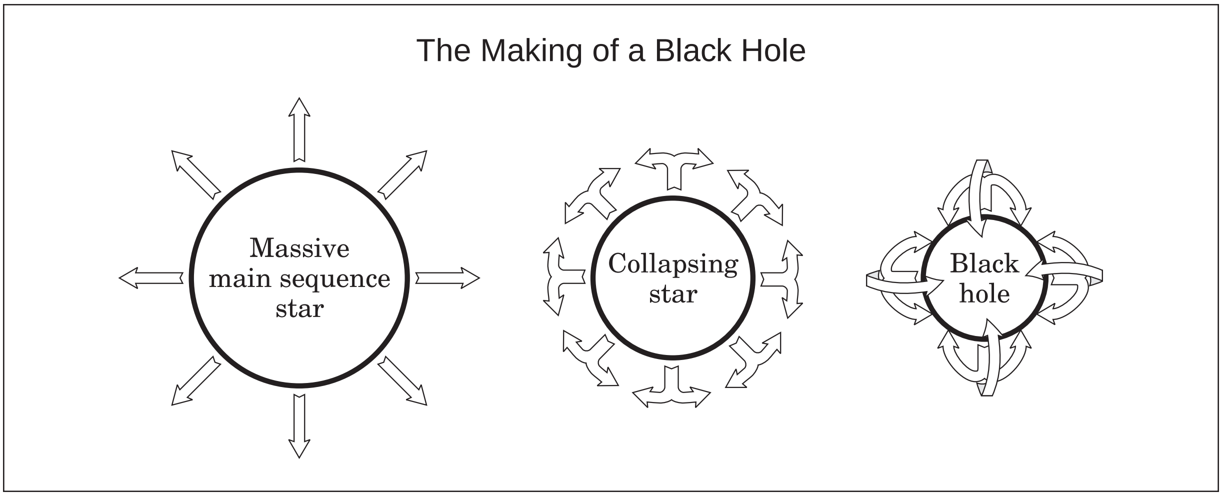
The heaviest element that can be synthesized inside a star is iron, which the star cannot burn. Once nothing is left in the star's core except iron, the star may follow two possible routes, depending upon its star's size. An average-size star such as Earth's Sun will expand and contract, periodically throwing off outer layers of stellar matter, a phenomenon called a nova. Eventually, this star will cool and shrink into a dense, planet-sized white or brown dwarf. Larger stars will undergo an incredibly fast collapse with the release of enormous amounts of matter and energy, enough to outshine an entire galaxy for several weeks. Such a collapsing massive star is called a supernova. After the supernova explosion, all that is left of the original star is an incredibly dense neutron star or a black hole.

An average-size main sequence star such as the Sun proceeds through this evolutionary development in approximately 10 billion years. The Sun is halfway through its life. More massive stars (for example, blue and red giants, and supergiants) proceed through this evolutionary development much more quickly, sometimes in as short a time as 10 million years.

During their lifetimes, all stars emit electromagnetic radiation, which behaves both as particles called photons and as waves consisting of crossed oscillating electric and magnetic fields. Electromagnetic radiation travels at the speed of light, approximately 300,000 kilometers per second. Various types of electromagnetic radiation exist, each type having a particular frequency and wavelength. The electromagnetic radiation spectrum ranges from low-frequency, long-wavelength radiations to high-frequency, short-wavelength radiations. The electromagnetic radiation spectrum includes, in order from low to high frequency, radio waves, television waves, microwaves, infrared radiation, visible light waves, ultraviolet radiation, X rays, and gamma rays. These radiations are produced and emitted from the thermonuclear fusion reactions inside stars. Certain stars exhibit intense bursts of various types of electromagnetic radiation during predictable time intervals. Many of these stars are called pulsating stars. They include Cepheid variables, RR Lyrae stars, binary or double-star systems, novae, supernovae, and the more exotic neutron stars and black holes.

Cepheid variables and the closely related RR Lyrae stars represent a class of stars that pulsate because they vary in their brightness, or luminosity, over time. These stars brighten, then dim in predictable, cyclic patterns that may last for several days for some stars or several weeks for others. Each such variable star has its own characteristic period—the time required for it to proceed through each cycle of brightening and dimming. Astronomers hypothesize that Cepheid variables are older main sequence stars that have used up their hydrogen fuel and are shifting to helium fusion. This shift causes the stars to be thermodynamically unstable, resulting in periodic outbursts of energy (that is, brightening).

Many main sequence stars exist as binaries, or double-star systems, in which two stars orbit around a common center of gravity. Many Sun-like stars exist in binary systems. As the two stars of a binary system orbit each other, occasionally one will pass in front of the other relative to Earth, thereby eclipsing the view of the companion star. When the eclipse occurs, combined light output from the two stars dims.



A massive star may end its life as a black hole: During its main sequence (left), radiation emits outward. As the core burns (center), the star begins to collapse in on itself. Finally (right), the increasing mass at the core is so great that gravity is extremely strong, preventing any radiation (including light) from escaping.

Some examples of double stars within Earth's galactic vicinity include Alcor-Mizar in the constellation Ursa Major (the Big Dipper) and Sirius A,B in the constellation Canis Major.

A nova is an eruptive, older star that ejects outer layers of gas as it slowly shrinks in size. One theory describing novae maintains that a nova could result from a binary system in which one star pulls matter away from the other, thereby creating instability in the thief star (the one receiving matter from the other star) and an occasional ejection of matter from this same star. Such a nova associated with a binary system would release large amounts of energy as well as matter. X-ray emissions have been detected from nova outbursts in binary systems.

A supernova—the catastrophic collapse and explosion of a massive star—releases in an instant more energy than all the stars in any given galaxy. This energy emission includes high-frequency electromagnetic radiation, such as X rays and gamma rays. Approximately 1×10^{50} ergs of energy are released in a supernova explosion. Following the supernova explosion, the remains of the collapsed star may be either a neutron star or a black hole, both of which are superdense objects that exert strong electromagnetic and gravitational fields. According to theory, whether a star becomes a neutron star or a black hole depends upon the mass of the pre-supernova star. A star having a mass at least 1.4 times greater than the mass of the Sun should become a neutron star following a supernova explosion. If a star has a mass at least ten times greater than the mass of the Sun, then the star should contract infinitely into a black hole after a supernova explosion. Numerous neutron stars have been located. Black holes are strongly inferred from images taken by the Hubble Space Telescope and Chandra X-Ray Observatory. Supermassive black holes are believed to power many galactic cores, including the Milky Way's.

Neutron stars form when the entire mass of the star collapses to perhaps a 16-kilometer radius, thereby producing incredible densities that crush matter. The density and pressure are so great that atoms cease to exist and protons and electrons fuse to produce neutrons. The neutron star, therefore, is one gigantic atomic nucleus composed solely of neutrons. If the neutron star is rotating, it is called a pulsar because it pulsates. Certain regions of the rotating star are hot spots of radiation emission, including X rays. As the pulsar rotates, the hot spots appear periodically with the associated radiation emission. Neutron stars are believed to exist at the centers of such supernova remnants as the Crab nebula in Taurus, Cassiopeia A, and Supernova 1987A. Neutron stars also appear to be associated with certain binary star systems. In such systems, one star is visible because of its light emission, but the neutron star companion is invisible because it does not emit visible light. The neutron star pulls matter away from its companion star, occasionally ejecting the matter with an intense burst of X rays.

A black hole is an even more extreme case of superdensity than a neutron star. A black hole is hypothesized to occur when a supernova remnant at least ten times more massive than the Sun collapses to infinite density, thereby creating a hole, or vortex, in space and time that permanently captures any matter or energy that enters the black hole's intense gravitational field. Not even light can escape from a black hole; however, black holes are intense emitters of X rays. Furthermore, black holes should cause intense X-ray outbursts from binary systems where a black hole draws matter away from a companion star.

Numerous stellar X-ray sources have been detected, primarily from the experiments of the American cosmic-ray astronomers Riccardo Giacconi and Herbert Friedman. These scientists and others launched X-ray detecting telescopes into orbit to locate X-ray stars that were undetectable from beneath the Earth's atmosphere. In 1962, Giacconi discovered the first extrasolar X-ray source, Scorpius X-1. Scorpius X-1 is one of several X-ray sources found within binary star systems. Several others include Centaurus X-3, Cygnus X-1, and Hercules X-1. Each of these binary star systems consists of two objects, one being a visible light-emitting star, the other being an invisible X-ray emitter. Scientists believe that the invisible X-ray emitter is pulling matter away from the visible companion star. There is considerable debate over whether the invisible X-ray emitters are neutron stars or black holes. In several cases, estimates of the invisible companion's mass within such binary systems favor a black-hole explanation.

Orbiting X-ray telescopes such as Uhuru (meaning "freedom" in Swahili) and the Einstein X-Ray Observatory first detected hundreds of stellar X-ray sources. One particular class of X-ray emitters releases enormous bursts of X rays over a period of minutes, hours, or days. Some of these X-ray bursters are associated with globular clusters, large accumulations of stars located around the halos of most galaxies, especially surrounding galactic centers. Other X-ray bursters appear to be associated with binary star systems containing neutron stars. In the latter case, the tremendous X-ray emissions most likely are resulting from the attraction of matter from the visible star to the invisible companion, which is probably a neutron star or a black hole. Other stellar X-ray sources include galactic nuclei and quasars. Galactic nuclei, the centers of galaxies, are compact clusters of billions of stars from which X rays emanate. The X-ray emissions could come from the combined stellar atmospheres of these billions of stars. An alternative hypothesis is that a supermassive black hole exists at the center of each galaxy. Because of the massive clustering of stars in galactic nuclei, this hypothesis seems plausible. There is strong evidence that black holes exist at the centers of the Milky Way galaxy

Quasars, quasi-stellar radio sources, are intense emitters of all forms of electromagnetic radiation, including visible light, radio waves, and X rays. Quasars are believed to be the oldest, most distant objects detected in the universe. They may represent the first major structures produced from the big bang explosion that created the universe, or they may represent vestiges of the big bang itself. One theory describing quasars maintains that they are supermassive black holes 1 billion times more massive than the Sun. Such black holes would be emitters of enormous quantities of energy and radiation.

X-ray sources appear to emanate predominantly from unstable processes that occur during certain phases of a star's life or during certain interactions between stars. The X-ray and other radiation emissions usually occur in intense bursts, which follow predictable, cyclical periods. All stars emit some radiation from thermonuclear fusion, but some (variables, binaries, novae, supernovae, bursters, neutron stars, and black holes) release large quantities at specific times. Research continued with even more sophisticated space-based observatories, principally the Hubble Space Telescope, Chandra X-Ray Observatory, and XMM-Newton. As a result of Hubble images, black holes moved from the realm of the theoretical to become accepted astrophysical objects to be studied across the electromagnetic spectrum.

APPLICATIONS

The discovery of high-level X-ray emitting objects in the universe helps astronomers to understand the nature and evolution of the universe, the evolution of stars, and the fate of Earth's star, the Sun. X rays and other high-frequency electromagnetic radiations are emitted by a variety of stellar sources (for example, stars, planets, galaxies, quasars, and interstellar hydrogen gas). These X-ray emissions are natural releases of energy from various chemical and physical processes.

Astronomers use a variety of sophisticated instruments to detect and map the sources of electromagnetic radiation emissions. Optical telescopes are used to measure the visible light-emitting objects, such as stars, planets, and galaxies. Radio telescopes are used to detect the radio emissions from the same objects. For X rays, however, detector telescopes must be launched above the Earth's atmosphere. Ground-based telescopes cannot detect X rays because X rays are absorbed by the atmosphere.

Optical and radio telescopes can detect variable stars, binary systems, novae, supernovae, neutron stars, galaxies, and quasars. However, X-ray detectors are needed above Earth's atmosphere to determine whether these objects release X rays and the quantity of radiation released. From the late 1950's through the early 1980's, astronomers at the Naval Research Laboratory in Washington, D.C.; at American Science and Engineering, Inc., in Cambridge, Massachusetts; and at the National Aeronautics and Space Administration (NASA), Goddard Space Flight Center in Greenbelt, Maryland, worked diligently at producing orbiting X-ray satellites. Among the principal scientists who were involved in this work were Giacconi and Herbert Gursky of American Science and Engineering, Inc., and Friedman of the Naval Research Laboratory. While initially attempting to map the location of radiation hot spots coming from the Sun and from around the Earth-Moon system (for example, the Van Allen radiation belts), they developed suborbital rocket-borne X-ray detectors to locate extraterrestrial X-ray sources. Friedman led the efforts with high-altitude rocket launches during which the X-ray detectors mapped X-ray emissions from the Sun.

In the summer of 1962, Giacconi, Gursky, and their colleagues launched a suborbital rocket above Earth's atmosphere. Aboard the rocket, a tiny X-ray detector located the first extrasolar X-ray source, Scorpius X-1, in the constellation Scorpius. Scientists have discovered that Scorpius X-1 is part of a binary system. Giacconi, Gursky, Friedman, and other astronomers followed this breakthrough with more suborbital X-ray telescope launches, locating many more X-ray emitters in the process.

In 1970, the first orbiting X-ray satellite, the Cosmic X-Ray Explorer, also called Uhuru, was launched from Kenya by Giacconi, Gursky, and other NASA astronomers and engineers. It detected and mapped the locations of hundreds of new X-ray sources during the early 1970's. It was followed in 1978 by the Einstein X-Ray Observatory, a more sensitive orbiting X-ray telescope that mapped even more X-ray sources outside the solar system during its three-year operational lifetime. Einstein was followed in 1999 by the Chandra X-Ray Observatory. The data collected from these and other astronomical projects fueled the theories of astrophysics and cosmology—specifically, the theories on stellar evolution and black holes. For the first time, many decades-old ideas could be tested. The existence of neutron stars was verified, X-ray emissions from the gas exchange between binary stars and novae could be substantiated, and considerable evidence supporting the existence of black holes within certain X-ray emitting binaries was obtained. X-ray astronomy caused a boom in the astronomy and astrophysics community, and it emerged as an important branch of astronomy and physics.

The existence of black holes had been hypothesized for more than two hundred years by numerous philosophers and astronomers. In the first half of the twentieth century, their existence was championed by the theoretical physicists Subrahmanyan Chandrasekhar, Fritz Zwicky, and John A. Wheeler. Black-hole physics was developed in meticulous detail by the theoretical physicists Stephen W. Hawking, Igor Novikov, and Kip Thorne, who explained the particle and energy emissions from these objects.

Black holes are the final state of stellar evolution for very massive stars. Most cosmological theories that describe the structure and evolution of the universe support the inevitable formation of black holes. These theories include Albert Einstein's theory of general relativity. The universe may be teeming with countless trillions of black holes, with several million in the Milky Way alone. Black holes may eventually accumulate all the mass of the universe within many trillions of years. This scenario is just one of many proposed end states of the universe, one in which all matter is absorbed into black holes. Even then, black holes gradually should evaporate their energy into space, leaving nothing but an energy-dominated universe with practically no matter left. By contrast, a more accepted end-state would have the universe continue to expand and suffer a cold, entropic death. The existence of X-ray sources in novae, binary systems, and galactic nuclei provides evidence supporting or rejecting theories describing the origin and evolution of the universe.

In the distant future, major X-ray bursters may be sources of considerable concern for human space travelers. High-frequency radiations, such as X rays and gamma rays, penetrate many materials and living tissue, producing mutations and sometimes death. X-ray astronomy will help astronomers to understand the universe and to live safely in space.

CONTEXT

X-ray astronomy has enabled scientists to see the universe from a different perspective. The fact that many stellar objects release intense bursts of X rays in predictable cycles has expanded the interpretation of how stars develop and interact with one another. Intense bursts of X rays are emitted from stars undergoing instabilities in their development or from violent interactions between pairs of stars, especially if one star in the pair is a neutron star or a black hole. Low quantities of X rays are emitted from a variety of stable objects, including the solar flares of the Sun and the atmospheres of gaseous planets such as Jupiter.

Among the most intense X-ray emitters are binary star systems, in which one member of the pair is a nova, neutron star, or black hole. Also, supernovae are intense X-ray emitters. Variable stars and novae by themselves are X-ray emitters, but not intense emitters. Large quantities of X rays are released from these objects in all directions when there is a violent transfer of matter away from one star onto the surface of another or when matter is ejected from a supernova at incredible speeds, temperatures, and pressures.

X-ray emissions from normal stars and from unstable interactions between stars usually are associated with intense magnetic fields on the star's surface. The interactions of matter with these fields cause the rearrangement of electrons within atoms with a corresponding release of energy such as X rays. The magnitude of the interaction determines the quantity of the X-ray burst.

With X-ray bursts coming from violent stellar interactions, it is interesting that quasars are such intense X-ray bursters, possibly reflecting a violent era of the early universe when matter and energy were ejected from black holes. The same situation may be seen in the future as matter gravitationally contracts and accretes into new black holes. X-ray bursts could represent telltale signs of black holes, which are scattered throughout the universe.

David Wason Hollar, Jr.

FURTHER READING

Fabian, A. C., K. A. Pounds, and R. D. Blandford. *Frontiers of X-ray Astronomy*. Cambridge, England: Cambridge University Press, 2004. Covers contemporary research with space-based X-ray telescopes. For the most serious astronomy reader or students of astrophysics.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory text covering the field of astronomy. Contains descriptions of astrophysical questions and their relationships.

Hawking, Stephen W. *A Brief History of Time:*

From the Big Bang to Black Holes. New York: Bantam Books, 1988. This enormously popular best seller, written by one of the leaders of twentieth century physics, is a clear, outstanding discussion of cosmology. Hawking describes the evolution of the universe, stellar evolution, and black holes.

Maccarone, Thomas, J. *From X-ray Binaries to Quasars: Black Holes on All Mass Scales*. New York: Kindle, 2006. Provides descriptions of high-energy processes that produce X-ray emissions. Describes the cosmological significance of quasars, black holes, and other high-energy objects.

Novikov, Igor, and Vitaly Kisin. *Black Holes and the Universe*. Cambridge, England: Cambridge University Press, 1995. A brief, clear presentation of big bang cosmology and black holes. Novikov, a leading Soviet theoretical physicist and black-hole researcher, describes the predicted properties of black holes and the role they should play in the evolution of the universe.

Schlegel, Eric M. *The Restless Universe: Understanding X-ray Astronomy in the Age of Chandra and Newton*. New York: Oxford University Press, 2002. Covers X-ray astronomy for early rocket-launched studies to the Chandra and XMM-Newton observatories. Explains the cosmological implications of current X-ray astronomy research. For the serious layperson.

Trumper, Joachin, and Gunther Hasinger. *The Universe in X Rays*. New York: Springer, 2008. Reviews the development of X-ray astronomy and its impact on advancements in astrophysics and cosmology. Covers ROSAT, RXTE, BeppoSax, Chandra, and XMM-Newton results.

Verschuur, Gerrit L. *The Invisible Universe: The Story of Radio Astronomy*. New York: Springer Praxis, 2006. Provides a history of developments in radio astronomy and along the way describes the discovery of pulsars, quasars, and radio galaxies. Suitable for a general science course in college as well as for astronomy majors as background information about the universe when examined in all segments of the electromagnetic spectrum.

Zeilik, Michael, and Stephen A. Gregory. *Introductory Astronomy and Astrophysics*. 4th ed. Fort Worth, Tex.: Saunders College Publishing, 1998. Excellent introductory astronomy book for undergraduate science majors. A well-written, extensively detailed survey of astronomy, with simple mathematics and physics. The numerous photographs, illustrations, and tables make this book an excellent reference for the serious amateur astronomer as well. Chapter 18, “Variable and Violent Stars,” discusses X-ray-emitting stars such as novae and bursters.

See also: Brown Dwarfs; Gamma-Ray Bursters; Main Sequence Stars; Nuclear Synthesis in Stars; Protostars; Pulsars; Red Dwarf Stars; Red Giant Stars; Stellar Evolution; Supernovae; Thermonuclear Reactions in Stars; White and Black Dwarfs.

Nuclear Synthesis in Stars

Category: The Stellar Context

With the exception of hydrogen, helium, and a small number of other light elements, the chemical elements found on Earth and throughout the universe were all formed in various stages of nuclear fusion within stellar interiors. Although these processes cannot be observed directly, models for element synthesis and the dissemination of that material to interstellar space help explain the relative amounts of elements observed in the Sun and other stars.

OVERVIEW

Nuclear synthesis is the process by which stars convert light elements, such as hydrogen and helium, into heavier elements, such as carbon, oxygen, magnesium, silicon, iron, and the rest of the stable elements found on Earth and throughout the cosmos. The process of nuclear synthesis requires conditions of high temperature and high density in order to proceed. Only under such extremes can positively charged nuclei, which must interact to form heavier species, overcome the enormous Coulomb repulsion that dominates under usual conditions. As a result, the dominant sites of nuclear synthesis in the present universe are in the cores of stars. Even at a temperature of 15 million kelvins and a tremendous density of 150,000 kilograms per cubic meter (typical of the environment at the center of the Sun), only rarely will an individual collision between nuclei successfully penetrate the Coulomb barrier and allow the nuclei to approach within the critical distance (10^{-15} meters, or 1 fermi, which is on the order of their radii) where the attractive strong nuclear force mediates their interaction. Outer layers of stars do not participate in the synthesis of the elements but are of great importance for the transport of energy generated by the nucleosynthesis processes taking place inside the core.

The first major nuclear reaction that occurs in stars is the conversion of four single protons to a helium 4 nucleus consisting of two protons and two neutrons (uncharged particles of roughly the same mass as a proton). This reaction requires several stages to complete. First two protons must fuse to form a nucleus of deuterium (also known as "heavy hydrogen" or hydrogen 2) consisting of one proton and one neutron. To maintain charge balance, a positively charged particle known as a positron, having the exact mass of an electron, is emitted along with an uncharged weakly interacting particle called a neutrino. The deuterium nucleus fuses with another free proton to form a helium 3 nucleus, which has two protons and a neutron. In the process, a high-energy gamma-ray photon is emitted. The entire sequence must be repeated by another group of protons to form a second helium 3 nucleus. The final stage of the reaction occurs when two helium 3 nuclei interact to form a single nucleus of helium 4, consisting of two protons and two neutrons. Two free protons are generated in this final step. A careful count of the particles involved indicates that the net reaction has resulted in a conversion of four free protons to a single nucleus of helium. The amount of energy released in each net conversion is exceedingly small; thus, many such reactions must be taking place within stars to produce enough energy to keep the temperature high enough to maintain the gas pressure required for them to hold up their massive outer layers against gravitational collapse.

A star like the Sun originally has sufficient supply of hydrogen to continue the synthesis of helium in its core for about 10 billion years. Subsequently, the lack of readily available fuel in the core of the star forces it to change its internal structure in an effort to find additional sources of energy. The predominantly helium core is at too low a temperature to initiate helium fusion, so it can only contract, releasing gravitational potential energy and growing hotter. This raises the temperature of the hydrogen immediately outside the core to the point that hydrogen fusion can occur in a shell around the core. The outer layers expand and cool, and the star swells to become a red giant. Meanwhile, the helium core continues to contract, and its temperature continues to rise. When the temperature reaches 100 million kelvins, the core begins to fuse helium into carbon.

Helium fusion is known as the triple-alpha process. The name derives from the interaction of three nuclei of helium 4, also known as alpha particles. Once again, a multistage reaction must occur. Two alpha particles fuse to form a nucleus of beryllium 8, which has four protons and four neutrons. In the process a gamma-ray photon is emitted. The beryllium 8 nucleus is unstable and will decay into alpha particles within 3×10^{-16} second. If, however, during this brief period, the beryllium 8 nucleus captures another alpha particle, the result is the formation of a stable carbon 12 nucleus, again with the emission of a gamma-ray photon.

The multiple stages required for both hydrogen and helium fusion, and the relative infrequency of successful fusion events at temperatures typical of the cores of stars, are responsible for moderating the consumption of nuclear fuel and the energy output of stars. If this moderation were not the case, the lifetimes of stars would be insufficient for the formation of planetary systems and for the evolution of life as it is known on Earth.

Once stars manage to build up carbon nuclei in their cores, the way is theoretically clear for stars to build up additional elements with relatively simple single-stage interactions. Alpha particle capture by carbon nuclei produces oxygen nuclei. Fusion of two carbon nuclei could form a magnesium nucleus, but it is more likely to produce a neon nucleus plus a helium nucleus. A great assortment of nuclear species can form, and with them, the variety of available pathways for further fusion increases rapidly. However, one additional requirement must be met. As nuclei with a greater number of positively charged protons are formed, the collisions required to fuse such nuclei must become increasingly energetic to overcome the proportionately larger Coulomb repulsion. Carbon 12, for example, which has six protons and six neutrons in its nucleus, does not fuse until a temperature of 600 million kelvins is achieved at the center of a star. Silicon 28 (with fourteen protons and fourteen neutrons in its nucleus) does not fuse until a temperature of 3 billion kelvins is attained.

Stars like the Sun never reach core temperatures high enough for significant carbon fusions, although small amounts of oxygen can form via alpha particle captures by carbon. Thus, the Sun will end its life with a slowly contracting carbon-oxygen core, while its outer layers are ejected eventually to form a "planetary nebula," a low-density shell of gas illuminated by the now-exposed high-temperature core. The core, which at this point is an extremely small, high-density sphere known as a white dwarf, will cool slowly and fade over many billions of years to become a black dwarf.

Elements heavier than carbon are formed by stars with a higher mass than the Sun. Stars with masses greater than about 4 solar masses but less than about 8 solar masses proceed through their evolution in much the same manner as lower-mass stars, only faster. The greater pressure exerted by the more massive overlying layers increases the temperature and hence the rates of nuclear reactions in the core; thus, such stars burn their available fuel more quickly than 1-solar-mass stars. A 5-solar-mass star, for example, will form a carbon-oxygen core within 200 million years. In these intermediate-mass stars, ignition of carbon may take place explosively, and the star may blow itself apart in a carbon-detonation supernova explosion; however, it is more likely that the star will have lost so much mass through strong stellar winds that its core will never reach carbon-detonation temperature.

In stars more massive than about 8 solar masses, the core temperature rises quickly enough to allow further nuclear burning to take place. Although the details are uncertain because of the complexity and variety of available nuclear burning pathways, the general pattern is clear. As high-mass stars burn heavier and heavier elements, each new nuclear species first ignites in the core, and then in a shell surrounding the core once the core has exhausted that nuclear fuel. Within 8 million years, a 25-solar-mass star builds up layer upon layer of heavy elements. Eventually, an inert iron core forms at the center of this star. Iron 56 nuclei can fuse into heavier nuclei only with the input of energy, so it is stable against further fusion. Instead, the iron core contracts rapidly in a vain attempt to liberate enough gravitational potential energy to hold up the outer layers of the star. Quickly, the core reaches temperatures at which the iron nuclei spontaneously disintegrate into individual alpha particles, which drains the core of energy. The star's outer layers collapse onto the high-temperature core, and a spectacular core-collapse supernova explosion results.

Alternative pathways for the formation of heavy elements exist, other than those caused by the fusion of alpha particles with nuclei such as carbon, magnesium, and silicon, or the fusion of these nuclei with themselves. During the evolution of stars, a fraction of all the usual fusion reactions produces free neutrons. These neutrons are captured subsequently by other nuclei, producing elements (or isotopes of elements) not synthesized by the usual charged nuclei process. A nucleus that captures an additional neutron is often unstable and will spontaneously "beta decay" after a time into an element with an additional proton. To maintain charge balance, the nucleus gives up an electron. The process of neutron capture and subsequent beta decay is known as the "s-process" ("s" for "slow," because of the relatively slow production rate of free neutrons in stellar interiors). Because the neutron is an uncharged particle, it need not overcome the repulsion of the positively charged nuclei in order to interact; thus, the s-process can proceed at relatively low temperatures. In the highly energetic environment of a supernova, many free neutrons are available. A nucleus may accept a large number of free neutrons before it has time to beta decay. This phenomenon, known as the “r-process” (“r” for “rapid,” because of the rapid addition of neutrons prior to the decay of the nucleus), accounts for the formation of elements heavier than iron. Eventually, a nucleus saturated with neutrons will decay, via the emission of alpha particles or electrons, into other nuclear species.

By a combination of charged-particle fusion reactions and neutron captures via the s- and r-processes, the stable elements in the periodic table are formed. It is important to note that a single star does not build up all the elements that are presently observed in its outer atmosphere. Indeed, several generations of stars must have been born, lived, and died in order for the elements now found in the Sun to have been formed. The elements that are created in previous generations of stars are dispersed and eventually mixed into the interstellar medium, the material between the stars from which future generations of stars ultimately form. Thus, the nuclear synthesis of elements is an ongoing process.

METHODS OF STUDY

Models of nuclear synthesis in stars are derived from a comparison of the observed abundances of the elements in stars, the Earth, the Moon, meteorites, and other solar-system objects against sophisticated computer models for the synthesis of the elements. Direct confirmation of the nuclear processes is not available, because the outer layers of stars prevent observation of the composition of stellar cores, where the products of element synthesis reside during a star’s lifetime. Further complications are introduced because the elemental abundances in relatively recently formed stars, such as the Sun, are the net result of nuclear synthesis products formed in many previous generations of stars with a wide range of stellar masses. Thus, comparisons of observation with synthesis models necessarily involve assumptions about the initial mass distribution of stars that formed in the Milky Way galaxy, the evolution of stars, the mass loss from the outer atmospheres of stars during their lifetimes, the nature of supernovae, mixing of elements in the interstellar medium, as well as subsequent processes, such as the effects of the impact of cosmic rays (high-energy charged particles) on nuclei in the interstellar medium. A self-consistent picture can be drawn in which the observations of elemental abundances agree with the model predictions reasonably well.

The most readily available test for nuclear synthesis models, and all of their associated assumptions, comes from a detailed inspection of the elements detected in the spectrum of light from the Sun. To a first approximation, the shape of the solar spectrum resembles that of the Planck curve of blackbody radiation—that is, the distribution of energy coming from a body in thermodynamic equilibrium. Superimposed upon this continuous spectrum are a multitude of dark lines that form when light is absorbed selectively by elements present in the solar photosphere. Patterns and strengths of absorption lines observed in terrestrial laboratories and calculated from detailed models of atomic-energy levels can then be compared to observed patterns and strengths of absorption lines found in the solar spectrum. In this way, a description of the relative abundances of elements in the Sun’s photosphere can be constructed. Since it is assumed the Sun was initially well-mixed, the photospheric abundances should be representative of the Sun’s original composition.

Other bright stars in the vicinity of the Sun have nearly the same relative abundances of elements in their atmospheres, because they formed at a similar epoch out of interstellar gas with much the same chemical composition. Some of these stars are hotter than the Sun; most are cooler. Patterns and strengths of absorption lines caused by the elements are affected greatly by the temperature of the stellar surface, providing a convenient check on models for line formation under different conditions. Stars that are much older than the Sun and its neighbors also play an important role in the comparison of models for nuclear synthesis in stars. Stars have been identified in the outer halo of the Milky Way galaxy that must have formed at very early times. From spectra of these old stars, astronomers obtain a picture of the production of elements at a time before nuclear syntheses from numerous generations of stars enriched the interstellar medium with heavier nuclei.

Abundances of elements formed in stars also can be obtained from chemical analysis of ancient rocks and meteorites found on the surface of Earth and the Moon. An inspection of such material provides a snapshot of the composition of the gas and dust from which the solar system formed about 4.5 billion years ago. Sophisticated laboratory equipment, such as particle accelerators and mass spectrometers, can be used to obtain very precise estimates of the relative abundances of the elements found in these rocks. Though such studies provide important clues, the interpretation is complicated by the chemical differentiation and segregation that must have taken place as a result of past geologic activity on Earth and the Moon. Samples of material obtained from the asteroid belt between Mars and Jupiter, as well as samples from the surfaces of comets, will be extremely important, as the asteroids and comets presumably have not suffered changes as great as those that have occurred on Earth and the Moon.

CONTEXT

Modern understanding of the formation of elements in the stars began with speculations for the source of energy required to sustain stars over their long lifetimes. Early ideas, which envisioned the Sun powered by gravitational potential energy liberated during a slow contraction, failed when it was demonstrated that the Earth itself was billions, rather than millions, of years old. In the early 1900's, it was realized that the extreme conditions of temperature and density at the centers of stars might be sufficient to set loose the energy that was locked away in the nuclei.

Physicists and astronomers in the 1920's and 1930's (such as Henry Norris Russell and Stanley Sir Arthur Eddington) outlined the basic processes by which the transformation of hydrogen nuclei into helium nuclei could lead to the energy generation required to power stars. Significant impetus for these ideas came from the use of the newly developed quantum mechanics, in particular by George Gamow, which demonstrated that a small fraction of the collisions between hydrogen nuclei in the core of the Sun could penetrate the repulsive Coulomb barrier. Further advances came in the 1950's, when a concerted attack on the problem of the formation of the elements was undertaken by the physicists Ernst Opik, Edwin Salpeter, Fred Hoyle, and others. Detailed data for the relative abundances of elements found in the Sun and on Earth available in the mid-1950's prompted an outline of the neutron capture processes jointly by Margaret and Geoffrey Burbidge, William A. Fowler, and Fred Hoyle, and independently by Alistair Cameron.

Dramatic confirmation of the theoretical ideas for nuclear burning in the late states of stellar evolution came with the supernova event of February, 1987 (Supernova 1987A), when a 20-solar-mass star in the nearby Large Magellanic Cloud galaxy exploded (about 160,000 light-years away). Supernova 1987A was the first nearby supernova to go off in modern times, and observations by astronomers from all over the world made it possible, for the first time, to obtain detailed spectra of the processed material as it emerged from this explosion. In the future, details of stellar evolution and the roles that stars of different mass play in the creation of the elements should become more clear. Studies of the oldest stars known, stars with substantially lower abundances of the heavy elements than are found in the Sun, will illuminate the nature of the first generation of objects that formed in the galaxy following the big bang.

Timothy C. Beers

FURTHER READING

Arnett, David. *Supernovae and Nucleosynthesis*. Princeton, N.J.: Princeton University Press, 1996. An undergraduate and first-year graduate study textbook that examines the evolution of stars. Also discusses galactic evolution, abundance of elements in stars and the solar system, and nucleosynthesis.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. A comprehensive summary of the knowledge of the solar system, written by scientists involved in the research. Contributions by Robert Noyes ("The Sun"), Paul Spudis ("The Moon"), and John Wood ("Meteorites") are particularly relevant to the topic of nuclear synthesis.

Chaisson, Eric. *Cosmic Dawn: The Origins of Matter and Life*. Boston: Little, Brown, 1981. An ambitious book that succeeds in presenting the "big picture" of the creation of the universe, the formation of galaxies, the nuclear synthesis of elements in stars, and the future of human culture in a consistent and comprehensible format. Chaisson's captivating writing style sustains the reader's interest throughout. Line drawings by Lola Chaisson enhance the discussion.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A very well-written college-level textbook for introductory astronomy courses. Contains a thorough description of the stages of stellar evolution, complete with transparent overlays, and an exceptionally lucid treatment of nuclear synthesis.

Clayton, Donald. *Handbook of Isotopes in the Cosmos: Hydrogen to Gallium*. Cambridge, England: Cambridge University Press, 2003. An attempt at explaining the origins of matter by investigating the abundance of various elements and isotopes. The author also explores what that abundance means through the theory of nucleosynthesis. Suitable for anyone familiar with astronomy, physics, and chemistry. Includes a glossary of technical terms.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Includes a good description of stellar evolution and nuclear synthesis.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook, thorough and well-written. Contains a good description of stellar evolution and nuclear synthesis.

Pagel, Bernard. *Nucleosynthesis and Chemical Evolution of Galaxies*. Cambridge, England: Cambridge University Press, 1997. An introductory textbook on galactic chemical evolution. Addresses thermonuclear reactions, chemical abundances, stellar evolution, and nucleosynthesis. For graduate students, scientists, and advanced undergraduates.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. This college textbook for introductory astronomy courses is divided into many short sections on specific topics. Presents a thorough discussion of stellar evolution and nuclear synthesis.

See also: Brown Dwarfs; Cosmology; Gamma-Ray Bursters; Hertzsprung-Russell Diagram; Main Sequence Stars; Novae, Bursters, and X-Ray Sources; Protostars; Pulsars; Red Dwarf Stars; Red Giant Stars; Stellar Evolution; Sunspots; Supernovae; Thermonuclear Reactions in Stars; White and Black Dwarfs.

O

Oort Cloud

Categories: Small Bodies; The Solar System as a Whole

The Oort Cloud marks the outer boundary of the solar system where the Sun's gravitational influence ends, and it answers the question: where do the long-period comets come from?

OVERVIEW

Comets are believed to be dirty snowballs. "Snow" means mostly water snow, but also frozen carbon dioxide, frozen carbon monoxide, frozen methane, frozen methanol, and perhaps others. "Dirty" refers to bits of carbon, silicate pebbles, and other materials mixing with theses ices. As a comet approaches the Sun, the ices sublime to vapor. The swiftly flowing vapors carry dust aloft, and the dust and vapor form a cloud called the "coma" around the cometary nucleus. The solar wind passing through the coma drags the vapors and dust out into a tail. If the comet comes close to the Earth and if the tail is spectacular, it can become visible to those on the surface of Earth.

Based on their orbital periods, comets are placed in one of two classes. Long-period comets have orbits that take more than two hundred years to complete, while short-period comets have periods of two hundred years or less. The orbits of short-period comets lie more or less in the ecliptic plane. The ecliptic plane is the plane of the Earth's orbit around the Sun. It is called the ecliptic because when the Moon is also in this plane, there might be an eclipse. The orbits of the other planets are also close to the ecliptic, but long-period comets come in from all angles with respect to the ecliptic, and from this astronomers infer that the source of long-period comets must be a spherical halo surrounding the solar system.

In 1950 a Dutch astronomer named Jan Hendrik Oort pondered the source of long-period comets. In examining the orbits of first-time, long-period comets, Oort found that on the average they came in from 44,000 astronomical unites (AU) from the Sun. (One AU is the distance from the Sun to the Earth, about 150 million kilometers.) Comets coming into the inner solar system for the first time have more volatile material on and just below the surface than repeat comets do, so first-time comets develop a coma farther away from the Sun and brighten more than repeat comets do. Oort proposed that there is a vast spherical cloud of comets extending 100,000 AU outward from the Sun with a concentration at 44,000 AU. This comet cloud has since been named the Oort Cloud.

Another peak in the aphelia (orbital points farthest from the Sun) of first-time comets occurs at 10,000 AU. This suggests that there is an inner Oort Cloud. This cloud is a wedge-shaped doughnut centered on the ecliptic plane. It begins about 1,000 AU out, where it blends outward from the Kuiper Belt and extends outward to 20,000 AU, where it shades into the spherical outer Oort Cloud. The inner Oort Cloud is estimated to contain five trillion comets totaling about 33 Earth masses. The outer Oort Cloud has about one trillion comets totaling about 7 Earth masses.

The concept has been modified. The outer limits of the cloud may extend outward 150,000 AU or 2.4 light-years, half way to the nearest stars. The Sun's gravity is so weak at these distances that comets there are easily affected by the gravitational pull of passing stars or by the tidal forces exerted by the Milky Way galaxy itself. The Sun orbits the center of the Milky Way galaxy, passing through the galaxy's spiral arms perhaps four times in 700 million years, and may come close to massive gas clouds called nebulae. The gravitational pull from these clouds may stir up the Oort Cloud of comets. Although they are generally too far apart, occasionally the comets themselves may come close enough to interact gravitationally. The results of any stirring of the Oort Cloud may be to randomize comet orbits so that they occupy a sphere, to strip some comets from the cloud and send them out into the space between stars, or to send them inward toward the Sun. Since comets from the Sun's comet cloud can escape into interstellar space, it may be that some of the comets in the Oort Cloud have drifted in from other stars. Such alien comets might be recognized by having different isotopic abundances of various elements than have home-grown comets. So far, no comets from other stars have been definitively identified.

Another possibility for stirring the Oort Cloud is that there might be a large planet or a small star far from the Sun but gravitationally bound to the Sun. Most stars have one or more companion stars gravitationally bound to them, so the Sun is in the minority if it has no companion. Using 2008 technology, an Earth-sized planet 300 AU from the Sun or a Jupiter-sized planet 600 AU out would remain undetected. If the Sun has a small brown dwarf companion, it would need to be even farther away to remain undetected. If such a solar companion existed and if it had a 30-million-year orbit, it would stir up the Oort Cloud every 30 million years as it passed through. As a result of such stirring, for two to three million years the number of comets striking the Earth would increase a few hundredfold. If large comets struck the Earth, a series of mass extinctions would result.

Periodic comet showers might also result from the Sun sinking below and then rising above the disk of the galaxy. The disk contains most of the stars, gas, and dust in the region of the Sun's orbit about the galactic center. As the planets revolve about the Sun, they move along with the Sun as it revolves about the center of the Milky Way. When the Sun is above the disk, the mass of the disk pulls the Sun downward toward the disk. The Sun will pass through and go below the disk while the gravity pulls the Sun back toward the disk. The Sun will stop its downward motion and begin to rise toward the disk. It will pass through the disk and rise above the disk. The cycle will repeat again and again. The motion is like that of a carousel horse moving up and down while the carousel carries the horse around and around. As the intensity and direction of the disk's gravitational attraction on the Oort Cloud changes, the cloud will be stirred up. This oscillation of the Sun through the galactic disk may have a period of 30 million years. If the Sun lacks a companion star, and so far there is no evidence for a companion, this will be the dominant external effect on the Oort Cloud.

If either the mass extinctions on Earth or crater age throughout geologic history were periodic, a nonterrestrial cause would be the most likely, since there are no known cycles on Earth with periods of tens of millions of years. Neither the crater age nor mass extinctions display a strongly periodic timing, however. A 30- to 35-million-year cycle would match some crater ages. The Chesapeake Bay Crater and the nearby Toms Canyon Crater are both dated at about 35 million years ago. The Popigai Crater in Siberia, Russia, is dated at 35.7 million years ago, and the Mistasin Crater in Labrador, Canada, is dated at 36 million years ago. It is believable that all four of the impacting bodies resulted from the same stirring of the Oort Cloud. The Chicxulub Crater in Yucatán, Mexico, is 65 million years old and marks the demise of the dinosaurs. Going back another 30 million years to 95 million years ago, however, there is no good candidate crater. The conclusion is that the Earth may be pummeled with comet showers from time to time, but there is no conclusive evidence of a regular cycle.

KNOWLEDGE GAINED

A possible Oort Cloud object was discovered in 2003 and named Sedna for the Inuit goddess of the sea. Its perihelion (orbital point closest to the Sun) is 76.16 AU, and its aphelion (orbital point farthest from the Sun) is estimated to be 892 AU, near the beginning of the inner Oort Cloud. It is 1,200 to 1,800 kilometers in diameter, goes around the Sun in about 11,000 years, rotates on its axis in about ten hours, and does not get warmer than 33 kelvins. A second possible member is designated 2000 CR_{105}. It is estimated to be 250 kilometers in diameter and has a 44.3-AU perihelion and a 397-AU aphelion.

Since it is not possible to observe any Oort Cloud object directly, with the possible exceptions of Sedna and 2000 CR_{105}, scientists must learn about them by studying comets, Oort Cloud objects that nature has delivered to our doorstep. Comets seem to be largely water ice along with carbon dioxide ice, carbon monoxide ice, methane ice, methanol ice, ammonia ice, amorphous (noncrystalline) carbon, silicates and other stony materials, sodium, carbonates, simple hydrocarbons, and clays. The Deep Impact mission found that Comet 9P/Tempel was covered by a dust layer tens of meters deep and that the comet's nucleus was 75 percent empty space, making it structurally weak. The minor constituents, especially, which differ from one comet to another, imply that comets formed at various distances from the Sun.

Astronomers have observed extended dust disks around many stars, including Vega, Fomalhaut, Beta Pictoris, and Epsilon Eridani. They correspond with what would be expected during the formation of planets and perhaps even during the initial ejection of icy asteroids into the outer reaches of their systems. The large number of objects in the Sun's early Kuiper Belt would have allowed them to smash into each other, making the larger ones larger and smashing smaller ones to dust. Radiation pressure would then drive the dust from the system, or gravity would pull it into the star, depending on the size of the dust particles. In any case, the system will be relatively free from dust once collisions cease.

CONTEXT

When astronomers first began finding planets about other stars by analyzing the star's light for minute Doppler shifts, they did not expect early results, but that is what they got. They found something no one imagined existed, namely Jupiter-sized planets very close to their parent stars, or "hot Jupiters" as they are often called. Hot Jupiters produce a large signal that was easy to detect. The immediate problem was that there should not have been enough material available to make a Jupiter that close to the parent star. Astronomers recognized that the current understanding of planetary formation might be insufficient. An alternative was that perhaps planets migrate away from the location where they were formed. Theoreticians worked on the problem and decided that they should have been including planetary migration all along. While the disk of gas, dust, and asteroids not yet incorporated into a planet is still present, planets can migrate.

Whether the planets migrate inward or outward depends on the stage of planet formation, the state of the disk near the planet, and the location of the planet in the disk. Jupiter migrated inward by an estimated 2 percent, while Saturn, Uranus, and Neptune migrated outward by an estimated 10 percent, 15 percent, and 30 percent, respectively. During the tens of millions of years of migration, they all flung icy asteroids inward and outward. Jupiter's strong gravity tossed them out into the Oort Cloud, while Saturn, Uranus, and especially Neptune launched them into the Kuiper Belt and perhaps also into the inner Oort Cloud. This explains conditions in our solar system and is consistent with what is known of extrasolar planetary systems and with observed dust disks around other stars.

Charles W. Rogers

FURTHER READING

Benningfield, Damond. "Where Do Comets Come From?" *Astronomy* 18, no. 9 (September, 1990): 29-36. Discusses the formation, structure, and dynamics of the Oort Cloud. For a general audience.

Davies, John. *Beyond Pluto: Exploring the Outer Limits of the Solar System*. Cambridge, England: Cambridge University Press, 2001. The book is most directly about the Kuiper Belt, but it also discusses the Oort Cloud. Excellent and easy-to-read.

Garlick, Mark A. *The Story of the Solar System*. Cambridge, England: Cambridge University Press, 2002. Beautifully and copiously illustrated, this volume addresses comets, the Kuiper Belt, and the Oort Cloud in sections at the end of Part 3. Accessible to nonspecialists.

Lin, Douglas N. C. "The Chaotic Genesis of Planets." *Scientific American* 278, no. 5 (May, 2008): 50-59. Gives the latest views on the formation of the solar system, including the formation of the Kuiper Belt and the Oort Cloud.

Malhotra, Renu. "Migrating Planets." *Scientific American* 281, no. 3 (September, 1999): 56-63. Describes how, during the formation of the solar system, Jupiter migrated inward, while Saturn, Uranus, and Neptune migrated outward by slinging icy bodies to the inner solar system, the Kuiper Belt, and the Oort Cloud.

See also: Ceres; Comet Halley; Comet Shoemaker-Levy 9; Comets; Dwarf Planets; Eris and Dysnomia; Impact Cratering; Kuiper Belt; Meteorites: Achondrites; Meteorites: Carbonaceous Chondrites; Meteorites: Chondrites; Meteorites: Nickel-Irons; Meteorites: Stony Irons; Meteoroids from the Moon and Mars; Meteors and Meteor Showers; Nemesis and Planet X; Pluto and Charon; Solar System: Origins.

Optical Astronomy

Category: Scientific Methods

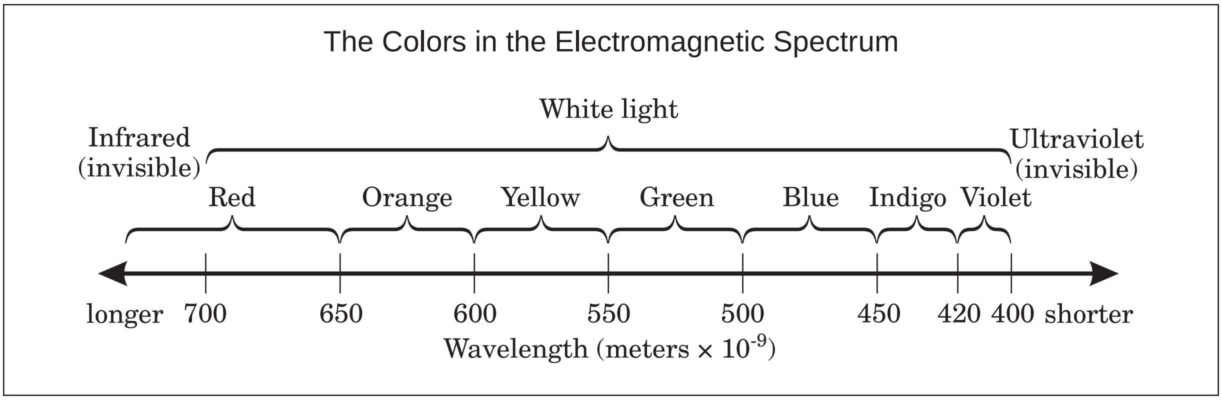
Optical astronomy encompasses all astronomical observations that can be made using visible light and adjacent wavelengths of the electromagnetic spectrum. This broad area includes highly detailed studies of intensity, polarization, spectral features, and other aspects of visible and near-visible light. Originally done first with the naked eye and later by looking through telescopes, optical astronomy evolved to the point where observations were made photographically during most of the twentieth century and digitally during the twenty-first.

OVERVIEW

Optical astronomy was the first form of astronomical observation. It relied on the human eye as an instrument capable of detecting light and discerning color and location. Today the term optical astronomy refers to observations made using visible and near-visible wavelengths of light, even if not made directly with the human eye.

Gamma, X, ultraviolet, visible, infrared, and radio waves all are forms of electromagnetic radiation. They differ only in wavelength, frequency, photon energy, and the instruments used to detect them. The wavelength range to which the human eye is sensitive, called visible light, spans only a small part of the entire electromagnetic spectrum, from violet light, at about 400 nanometers (10^{-9}, or one-billionth of a meter), to red light, at about 700 nanometers. This is the same wavelength range that the Sun emits most intensely, which is no coincidence; human eyes evolved to react to the dominant wavelength range of sunlight. Furthermore, the Earth's atmosphere blocks much of the electromagnetic spectrum from reaching the ground. One of the two major wavelength ranges that can penetrate the atmosphere and reach the ground produces the so-called optical window; the atmosphere is almost completely transparent between about 300 nanometers in the ultraviolet and 1,000 nanometers in the infrared, and partially transparent out to about 10,000 nanometers. (The other range is the "radio window," between about 1 centimeter and 10 meters.) Today, optical astronomy is considered as spanning the range from approximately 300 nanometers (the near-ultraviolet) to approximately 1,100 nanometers (the near-infrared), based on the optical window and the available detectors and instruments.

The optical systems that astronomers use are designed to deliver a beam of light to the surface of a particular detector or sensor, where it causes a change in the electrical or chemical state of the sensor material. This change is then recorded in some fashion and analyzed to produce information about the object that emitted the light. The two basic types of optical telescopes, refractors and reflectors, make use of two properties of light: it can be refracted (bent) when it passes through a transparent medium (a lens), and it can be reflected at the surface of a material (a mirror). The first telescopes developed in the early 1600's were refractors using lenses. However, reflectors using mirrors are more versatile, allowing a variety of optical paths to focus light. (Because reflection is the same at all wavelengths without regard to color, the same basic layout of visible light



The human eye perceives discrete bands of the electromagnetic spectrum as particular colors. Radiation of longer or shorter wavelengths is invisible, whereas white light contains all visible wavelengths. Wavelength is measured here in meters × 10^{-9}.

reflectors is used in most nonvisible-radiation detectors as well, including radio, infrared, ultraviolet, and X-ray telescopes. The principal differences have to do with the materials used for the mirrors and the detectors employed for different parts of the spectrum.) Some optical telescopes actually use a combination of mirrors and lenses to focus light into an image. In spectroscopy, additional optics are needed to take light from a chosen object and disperse it by wavelength into a spectrum. This is done using a prism (a wedge-shaped piece of glass) or a diffraction grating (which has lots of fine grooves scribed into the surface of some material). In either case, different wavelengths are bent different amounts. The spread-out spectrum can then be focused on a photographic or electronic detector.

The human eye was the first detector used behind the telescope. The development of photography provided a means of preserving telescopic images for later study. Astronomical photographs usually were made on glass plates coated with photographic emulsion because of the stability of glass, but plastic-based films were sometimes employed where a series of pictures were to be taken and there was no convenient way to change plates. Until the 1980's, photography was the method of choice for most observations. Later, digital electronic detectors came into wide use because they could record images more quickly than photography could. The charge-coupled device (CCD) is an electronic analog to the retina in the eye. A CCD is an array (like squares on graph paper) of little light sensors called picture elements (or pixels) that are read by a computer. CCDs provide a direct measure of the intensity of the visible light (or other form of radiation) striking each pixel and produce images that are stored in and can be directly manipulated by computers.

Until 1990, optical astronomy was largely conducted from the surface of the Earth, since space-based astronomy concentrated on regions of the electromagnetic spectrum that could not be detected from the ground. Although a number of small optical telescopes were carried on suborbital and short orbital platforms, optical astronomy did not truly move into space until 1990, when the Hubble Space Telescope (HST) was launched into Earth orbit. Above the atmosphere, it can obtain much sharper images than from the ground, where atmospheric turbulence always blurs images (a phenomenon astronomers refer to as "seeing") at least to some extent. The HST has provided a wealth of high-resolution images after initial post-launch optical problems were corrected.

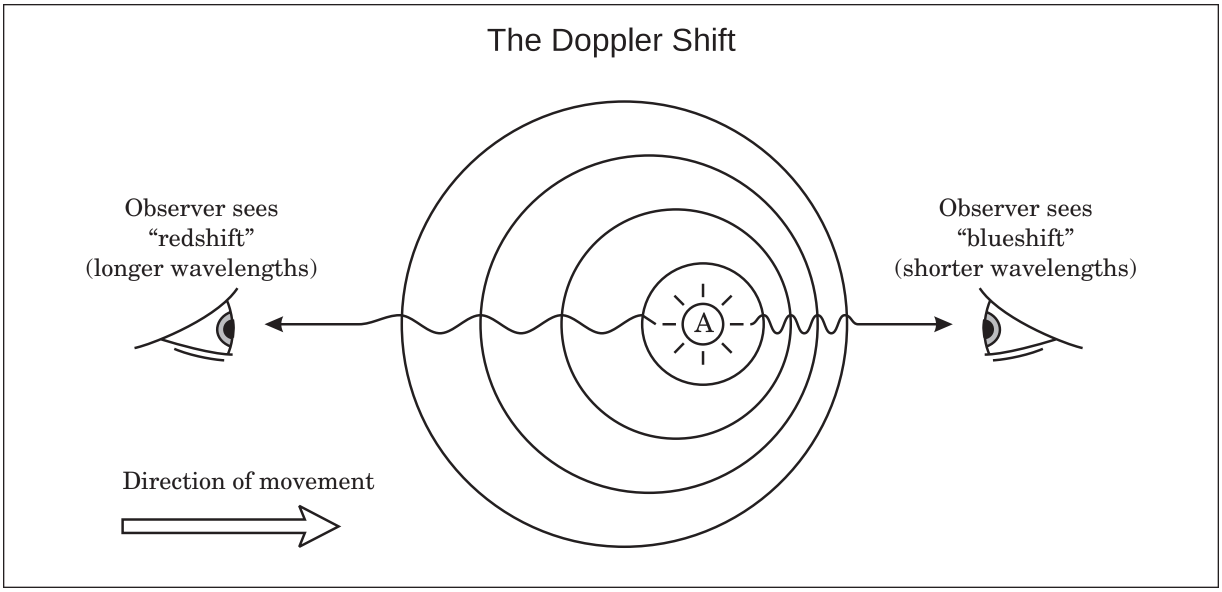
APPLICATIONS

Optical astronomy observations can be divided into several techniques: imagery, photometry, spectroscopy, and polarimetry. These, in turn, are applied to studying various sorts of objects.

Imagery provides "pictures" of objects, showing their locations, shapes, and other features. Some of the first measurements in optical astronomy were made with micrometers attached to telescope eyepieces so visual observers could measure accurate positions for constructing star charts. (Earlier charts were made using naked-eye pointing devices.) Photography and electronic imaging provide a means of recording images that can be studied and measured later and that reveal details that cannot be seen with the eye peering through a telescope.

Photometry, literally "light measure" or measurements of brightness, also was performed initially with the eye but was highly subjective; it was difficult to form a standard of brightness based on human perception. Then photographic and electronic systems became available. With photographic emulsions, brightness could be measured objectively, by the degree of exposure recorded. With solid-state photon counters and other electronic sensors, the instantaneous brightness of an object can be measured at any time, and quantitative comparisons can be made with measurements taken on different nights or among different objects. Brightness variations too quick or subtle for the eye to detect also became measurable. Both photographic and electronic photometry measurements can be made using all available light or limited to some particular wavelength range using specific colored filters.

Spectroscopy involves dispersing the light by wavelength using a prism or diffraction grating. It allows astronomers to determine a broad range of attributes of an object. Transparent gases produce "emission line" and "absorption line" spectra with distinct bright emission or dark absorption lines at particular wavelengths, while solids, liquids, and opaque gases produce continuous emission or reflection spectra (all wavelengths present with the intensity varying continuously by wavelength). The pattern of emission or absorption lines is determined by the chemical elements that are present in the transparent gas as well as its temperature. The variation of intensity with wavelength in a reflection spectrum from a solid or liquid can be compared to reflection spectra of laboratory samples to determine the probable surface material of the source. The variation of intensity with wavelength in a continuous emission spectrum closely follows a theoretical curve called the Planck distribution for thermal radiators; the wavelength at which the intensity



In the phenomenon known as the Doppler shift, light waves appear bluer (because they are shorter) when the source is moving toward the observer and redder (because they are longer) when the source is moving away from the observer. Above, "A" is a source of light, such as the Andromeda nebula.

reaches a maximum indicates the surface temperature of the source. Motion of the source toward or away from the observer (or detection device) along the line of sight shifts the entire spectrum and all its features to shorter or longer wavelengths as a result of the phenomenon known as the Doppler effect. If the source is moving toward the observer, all the wavelengths it emits or reflects are compressed, and the spectrum is shifted toward shorter (bluer) wavelengths; this is called a blueshift. If the source is moving away from the observer, all the wavelengths it emits or reflects are stretched, and the spectrum is shifted to longer (redder) wavelengths; this is called a redshift. The amount of the spectral shift depends on the speed of approach or recession. The intensity of magnetic fields can be measured by Zeeman splitting, in which pairs or triplets of spectral lines appear where a single line would be expected.

Polarimetry measures the linear or circular polarization of light. Polarization indicates certain conditions at the source of the emission or along the light path. The wave model treats light as associated electric and magnetic fields (hence the name electromagnetic radiation) that oscillate in two perpendicular planes as light propagates. It may be unpolarized (with no particular orientation to the planes of oscillation), linearly polarized (with the two planes of oscillation keeping the same orientation), or right or left circularly polarized (with the two planes of oscillation rotating completely around the light path through each wave cycle). Most light is emitted unpolarized but can become polarized by reflecting off a surface or being scattered by dust particles (which polarizes sunlight in the atmosphere). Light can also be polarized when emitted in a strong magnetic field (such as produces the Zeeman effect in spectra), or when emitted by electrons trapped in a magnetic field (synchrotron radiation). Most polarimetry is conducted with detectors that use the property of birefringence, which retards light waves vibrating at right angles to other light waves. Such detectors employ certain crystals that exhibit birefringence, or Pockels cells, which become birefringent in response to an electric current.

CONTEXT

After millennia of having nothing more to use than the unaided eye, optical astronomy got a real boost in the 1600's when Galileo manufactured his own refracting telescopes and used them to study the sky. His discovery of mountains and valleys on Earth's moon, spots on the Sun, moons orbiting Jupiter, and many stars too faint to be seen without a telescope (including those that constitute the hazy white band of light called the Milky Way) spurred further telescopic observations and more discoveries. Sir Isaac Newton designed the first reflecting telescope, although problems in fabricating and silvering the mirrors caused reflectors to lag behind refractors in size and capability for many years. (Today, all large optical telescopes that are built are reflectors.)

Even after the introduction of the telescope, for 250 years more optical astronomy relied on the human eye as the only detector and the brain and hand as the only interpreters and recorders of data. The advent of photography sparked a radical change in astronomy. The development of dry photographic plates in the late 1800's allowed long-duration exposures to be made for imaging faint, extended, diffuse objects whose central bodies astronomers had been able to view only dimly by eye. Photography also recorded images objectively for later study by any who were interested. The advent of electronic detectors in the 1980's made recording light even more efficient. Improved quantum efficiency of electronic detectors has given new life to old observatories, allowing them to record much more of the light (up to 60 to 70 percent now versus 1 percent photographically in the 1950's) delivered by their telescopes.

The application of spectroscopy to optical astronomy had a major impact in that the chemistry of astronomical sources could be assayed by the distinct patterns of bright or dark lines in the spectra of the light they emitted or reflected. In the early 1800's, Joseph von Fraunhofer discovered that the solar spectrum had distinct black lines in it, leading to the discovery that each chemical element emits or absorbs distinct wavelengths of light, thus providing a powerful tool for analyzing chemical compositions. This led to the discovery of helium in the solar atmosphere before its discovery on Earth. Perhaps the most striking contribution of optical spectroscopy was the discovery by Edwin Powell Hubble of the expansion of the universe based on how most galaxy spectra were shifted toward longer wavelengths (the red end of the visible spectrum). He found that spectral redshifts of galaxies were correlated with their distances, the farther galaxies having larger redshifts. He deduced that this redshifting was due to galaxies moving away from each other with speeds proportional to their distances apart; in other words, space is expanding.

Optical astronomy also provided the first observational confirmation of a prediction of general relativity with measurements (during a solar eclipse) showing that the Sun slightly bends starlight passing close by its edge. This effect has now been seen even more dramatically with the discovery of gravitational lensing by galaxies that bend the light of more distant objects to create multiple images.

The observational techniques of optical astronomy can be applied to study a wide range of astronomical objects—such as planets, stars, nebulae, and galaxies. Optical astronomy has served as the foundation for all the other branches of observational astronomy that deal with longer or shorter wavelength ranges, and it continues to interact with them as astronomers try to identify optical counterparts of new objects found in the radio, infrared, ultraviolet, X-ray, and gamma-ray spectral regions. Major advances have been made in observing all parts of the electromagnetic spectrum, because as new technology makes progress possible in one wavelength range, there is increased effort to bring the others up to par with it. Optical astronomy continues to serve as the linchpin for astronomy as a whole, and the future for optical astronomy remains bright.

Dave Dooling

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A well-written college-level textbook for introductory astronomy courses. Has three chapters on light and telescopes.

Field, George. *The Space Telescope*. Chicago: Contemporary Books, 1989. Description of the origins of the Hubble Space Telescope and the problems of ground-based and space-based astronomy. Includes descriptions of how optics and detectors work.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Has two chapters on light and telescopes.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. Thorough and well-written college-level introductory astronomy textbook. Includes two chapters on light and telescopes.

Hirsh, Richard F. *Glimpsing an Invisible Universe: The Emergence of X-Ray Astronomy*. New York: Cambridge University Press, 1983. Early history of X-ray astronomy through the High-Energy Astronomical Observatories. Written for the informed reader.

Kitchin, C. R. *Astrophysical Techniques*. 4th ed. Philadelphia: Adam Hilger, 2004. Discusses various methods used to study astronomy, along with instruments. A good reference for astronomy students.

Krisciunas, Kevin. *Astronomical Centers of the World*. Cambridge, England: Cambridge University Press, 1988. A survey of the major ground-based observatories and a description of their capabilities.

McLean, Ian S. *Electronic Imaging in Astronomy: Detectors and Instrumentation*. New York: Springer, 2008. Covers advancements and developments in astronomical detectors since 1970. Written by an expert in the field for a wide audience.

Moore, Patrick. *Astronomical Telescopes and Observatories for Amateurs*. New York: W. W. Norton, 1973. Introduction for the amateur astronomer to the basics of optical astronomy. Most details apply, though, to large-scale observatories.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A thorough college textbook for introductory astronomy courses, divided into many short sections on specific topics. Offers several sections dealing with light, telescopes, and observing techniques.

Tucker, Wallace. *The Cosmic Inquirers: Modern Telescopes and Their Makers*. Cambridge, Mass.: Harvard University Press, 1986. Literate, well-written descriptions of modern observatories and the scientific questions that led to their construction.

Wall, J. V., ed. *Optics in Astronomy*. New York: Cambridge University Press, 1993. A collection of articles written by experts throughout the world. Discusses the fundamental role optics plays in modern astronomy.

See also: Archaeoastronomy; Coordinate Systems; Earth System Science; Gravity Measurement; Hertzsprung-Russell Diagram; Infrared Astronomy; Neutrino Astronomy; Radio Astronomy; Telescopes: Ground-Based; Telescopes: Space-Based; Ultraviolet Astronomy; X-Ray and Gamma-Ray Astronomy.

P

Planetary Atmospheres

Category: Planets and Planetology

The gaseous envelopes that surround the planets are known as their atmospheres. Scientists believe that planetary atmospheres developed in one of two ways: either when the solar system as a whole formed, or later, from materials released by the planets well after their own formation in a process called outgassing. One task of continuing spacecraft investigations is to seek answers to lingering questions about planetary atmospheres.

OVERVIEW

Atmospheres are an attribute shared by seven of the eight planets in the solar system. (Since 2006 Pluto has been classified by the International Astronomical Union not as a planet but as a dwarf planet.) Crater-covered Mercury is the only planet that does not have a gaseous envelope. Even Titan, a large satellite of Saturn, possesses a significant atmosphere. Nevertheless, these atmospheres differ considerably in chemical composition, structure, and density, and comparative planetologists' study of these differences is key to an understanding of Earth's atmosphere and its future.

Planetary scientists believe the planets acquired their atmospheres in one of two ways. In some cases, the atmosphere formed with the planet around the Sun out of the solar nebula around 4.5 billion years ago. In such cases, a planet is said to have a primordial, or captured, atmosphere. In other examples, an atmosphere appears to have been created from gaseous material released from within the planet after it formed. These are called secondary, or outgassed, atmospheres.

Atmospheres of the Jovian planets (Jupiter, Saturn, Uranus, and Neptune) seem to be of the primordial type. Like the Sun itself, they are composed largely of hydrogen and helium. On the other hand, three of the terrestrial planets (Venus, Earth, and Mars) may have secondary rather than primary atmospheres. Characteristics and compositions of these atmospheres vary considerably. The dwarf planet Pluto, which is neither Jovian nor terrestrial in nature, does have a very thin primordial atmosphere that freezes on its surface during the part of Pluto's highly eccentric orbit when the dwarf planet is far from the Sun.

An important determinant of a planet's atmospheric composition and density is its ability to retain atoms and molecules of gases with its own gravity. According to kinetic molecular theory, gases are composed of discrete atoms and molecules traveling in random directions at a wide variety of different speeds. Collisions between these atoms and molecules occur, tranferring energy and altering the speeds of individual atoms or molecules. The distribution of those molecular speeds for the gas as a whole is governed by the temperature of the atmosphere according to the physics of Maxwell-Boltzmann statistics. Should the speed of any individual atom or molecule at some instant exceed the escape velocity of the planet, that atom or molecule will leak away into space. The root-mean-square speed of molecular motion in the gas is a function of the composition and temperature of the gases, while a planet's escape velocity is a function only of its mass.

The ability of a molecule to escape the atmosphere is also limited by the likelihood of collision with other molecules. In the lower atmosphere, this likelihood is so high as to be a virtual certainty. However, with increasing distance above the surface, the density of the atmosphere diminishes until a point is reached where a molecule moving upward and away from the planet has a very high probability of not encountering any other molecule. At that point, if its velocity exceeds the escape velocity, it will leave the planet. This region of a planet's atmosphere is called the exosphere.

Although by far the greatest amount of material in the early solar system was hydrogen and helium, today these gases are relatively rare in the atmospheres of the terrestrial planets. In part, this results from the fact that proximity to the Sun raises atmospheric temperatures of these inner planets to the point that molecular activity of hydrogen and helium atoms exceeds the low escape velocities of small planets. Some theorists argue, however, that as the solar system formed, most of the interplanetary hydrogen and helium in the inner solar system was pushed away from the region in which the terrestrial planets would later form by the pressure of radiant energy coming from the young Sun as soon as it began to shine.

Whichever theory is true, Earth's atmosphere today is approximately 78 percent nitrogen and 21 percent oxygen. Of the remaining 1 percent, most is argon and only 0.033 percent is carbon dioxide. At sea level, the atmosphere has a density of 0.001 gram per cubic centimeter. This atmosphere is divided into layers according to the prevailing characteristics at various altitudes. The lowest layer is the troposphere. It contains 90 percent of the atmosphere's mass and almost all the water vapor and dust. What sets the troposphere apart is the process of convection. Warm air currents rise and cool air currents sink. This distributes the heat from the Sun, and establishes the dynamics that create weather. The upper limit of this layer, called the tropopause, is defined by a temperature of 333 kelvins.

Above the troposphere is the stratosphere, which extends to an altitude of 55 kilometers. In the stratosphere temperature steadily increases with altitude. Within this layer, between 25 and 45 kilometers high, lies a sublayer in which ozone molecules occurs in a number density of ten parts per million. This trace concentration is enough to absorb ultraviolet radiation very efficiently, thereby protecting life. The ozone layer also prevents photochemical reactions that would deplete the planet of its water.

The stratosphere ends and the mesosphere begins at the point where temperature reaches about 283 kelvins and begins dropping again. The mesosphere continues to an altitude of 80 kilometers, where the temperature is about 163 kelvins. Above the mesosphere lies the ionosphere, so named because a significant number of the atoms of gas in this layer have been split into positively charged ions and negatively charged electrons by the Sun's intense X-ray and ultraviolet radiation. The Earth's exosphere—the layer that marks the transition to outer space—begins at 400 kilometers and extends for thousands of kilometers. The determination by the National Aeronautics and Space Administration (NASA) that the atmosphere begins at an altitude of 400,000 feet, the point where reentry of spacecraft begins, is therefore somewhat arbitrary.

Presence of free oxygen in abundance is unique to the atmosphere of Earth. Whether Earth's original atmosphere was of the primordial or outgassed type, free oxygen would have resulted in part from the photochemical dissociation of water vapor (until the development of an ozone layer) and then from photosynthesis, the process whereby early microorganisms such as cyanobacteria and later plants generate energy-supplying carbohydrates from water vapor and carbon dioxide in the presence of sunlight and chlorophyll. There is fossil evidence of photosynthetic plants existing on Earth going back at least 2 billion years ago. It is probably not coincidental that rocks 1 to 1.5 billion years old are the earliest to show strong evidence of having formed in an oxygen-rich atmosphere.

The atmosphere of Venus is ninety times more massive than Earth's and is composed of 97 percent carbon dioxide. The remaining 3 percent is almost entirely nitrogen, with trace amounts of water vapor, sulfur dioxide, and other gases. The atmosphere of Venus is very hot, resulting from the fact that the dense envelope of carbon dioxide is an excellent absorber of infrared energy and very effectively traps heat radiated by the surface. This phenomenon, called the greenhouse effect, causes the temperature of the atmosphere near the surface to reach 733 kelvins. Because of this high temperature, any water on the planet should exist as water vapor in the atmosphere. Its very low abundance (150 parts per million) presented a major puzzle for planetologists. Apparently, the high atmospheric temperature causes water vapor to rise to a great height, where it is then subjected to photochemical dissociation by strong ultraviolet radiation.

Planetary Atmospheres: Comparative Data

	Mercury	*Venus*	*Earth*	*Mars*
Surface pressure	~10^{-15} bars	92 bars	1.014 bars	6.1-9 millibars
Surface density (kg/m^3)	—	~65	1.217	~0.020
Avg. temperature (kelvins)	440	737	288	~210
Scale height (km)	—	15.9	8.5	11.1
Wind speeds (m/sec)	—	0.3-1.0	up to 100	2-30
Composition*				
Ammonia	—	—	—	—
Argon	tr?	70 ppm	9,430 ppm	1.6%
Carbon dioxide	tr?	96.5%	350 ppm	95.32%
Carbon monoxide	—	17 ppm	—	—
Ethane	—	—	—	—
Helium	6%	12 ppm	5.24 ppm	—
Hydrogen	22%	—	0.55 ppm	—
Hydrogen chloride	—	tr	—	—
Hydrogen deuteride	—	—	—	0.85 ppm
Hydrogen fluoride	—	tr	—	—
Krypton	tr?	—	1.14 ppm	0.3 ppm
Methane	—	—	—	1.7 ppm
Neon	tr?	7 ppm	18.18 ppm	2.5 ppm
Nitrogen	tr?	3.5%	78.084%	2.7%
Nitrogen oxide	—	—	—	100 ppm
Oxygen	42%	—	20.946	0.13%
Potassium	0.5%	—	—	—
Sodium	29%	—	—	—
Sulfur dioxide	—	150 ppm	—	—
Water	tr?	20 ppm	1%	210 ppm
Xenon	tr?	—	0.08 ppm	—

* Composition: % = percentages; ppm = parts per million; tr = trace amounts.

The dense atmosphere of Venus is efficient at circulating heat. The day and night temperatures of the troposphere do not differ appreciably. Huge, slow-moving convective air currents called Hadley cells span the planet's temperate latitudes from the poles to the equator and are not affected much by the planet's slow rotation. At elevations above 50 kilometers, where the atmospheric density is much lower than at the surface (actually about the same as the density of Earth's lower atmosphere), a strong westward current prevails over all but the polar regions and produces 360-kilometer-per-hour winds that whip the clouds into enormous swirling patterns.

Astronomers have long realized that the brilliance of Venus in the night sky is caused by the high albedo of the planet's perpetual cloud cover. Seventy-five percent of the sunlight reaching Venus bounces off the tops of a nearly featureless and visually impenetrable deck of sulfuric acid-laden clouds suspended between 47 and 70 kilometers above the surface. The clouds result primarily from photochemical reactions driven by strong ultraviolet energy from the Sun; hence, they have more in common with Earth's ozone layer than with Earth's clouds of condensed water vapor.

Mars has an atmosphere that is only 1/150 the mass of Earth, but carbon dioxide is the dominant component, accounting for 95.3 percent of its mass. Nitrogen accounts for an additional 2.7 percent and argon 1.6 percent of the

Planetary Atmospheres: Comparative Data *(continued)*

	Jupiter	*Saturn*	*Uranus*	*Neptune*
Surface pressure	>100 bars	>100 bars	>100 bars	>100 bars
Surface density (kg/m^3)	~0.16	~0.19	~0.42	~0.45
Avg. temperature (Kelvin)	~129	~97	~58	~58
Scale height (kilometers)	27	59.5	27.7	~20
Wind speeds (meters/second)	up to 150	up to 400	up to 200	up to 200
Composition*				
Ammonia	260 ppm	125 ppm	—	—
Argon	—	—	—	—
Carbon dioxide	—	—	—	—
Carbon monoxide	—	—	—	—
Ethane	5.8 ppm	7ppm	—	1.5 ppm
Helium	10.2%	3.25%	15.2%	19%
Hydrogen	89.8%	96.3%	82.5%	80%
Hydrogen chloride	—	—	—	—
Hydrogen deuteride	28 ppm	110 ppm	~148 ppm	192 ppm
Hydrogen fluoride	—	—	—	—
Krypton	—	—	—	—
Methane	3,000 ppm	4,500 ppm	2.3%	1.5%
Neon	—	—	—	—
Nitrogen	—	—	—	—
Nitrogen oxide	—	—	—	—
Oxygen	—	—	—	—
Potassium	—	—	—	—
Sodium	—	—	—	—
Sulfur dioxide	—	—	—	—
Water	—	~4 ppm	—	—
Xenon	—	—	—	—

Source: Data are from the National Space Science Data Center, NASA/Goddard Space Flight Center.

total, with small fractions of a percentage each of oxygen, carbon monoxide, and water vapor. Apart from the occasional planet-enveloping dust storms, Martian skies are generally clear, but several types of water vapor clouds do form in localized areas. Clouds of carbon dioxide are also noted occasionally.

Like Venus, Mars has a very low water vapor content in its atmosphere, but for a different reason. The very low pressure and temperature of the Martian atmosphere are not conducive to holding moisture at any elevation. In fact, a water vapor content of a few hundredths of a percent constitutes high relative humidity on Mars. This results in the appearance of water vapor clouds and fogs at lower elevations.

Mars exhibits seasonal wind patterns similar to those of Earth, with prevailing westerlies, high-altitude jet streams, and cyclonic disturbances. Strong localized winds at the surface are not uncommon, sometimes reaching 150 kilometers per hour. These are believed to be associated with powerful convection currents generated by the rapid heating of the surface during the daytime and are similar to dust devils, an atmospheric phenomenon common in the desert regions of the Earth. Dust devils on Mars were imaged from the surface by the Spirit and Opportunity Mars Exploration Rovers; those dust devils had the coincidental benefit of blowing dust off these rovers' solar panels, thereby increasing the electrical power available to the rovers. Once such a wind pattern forms, it may lift Martian dust high into the atmosphere, where the dust grains have a self-propagating effect, retaining heat and intensifying the wind disturbance. Periodically, such dust devils grow into planet-encompassing storms that may last for months. The Martian atmosphere also exhibits a strong seasonal wind flow caused by the exchange of carbon dioxide between the atmosphere and the polar caps.

Atmospheres of the "gas giant" Jovian planets contrast sharply with those of Venus, Earth, and Mars. Jupiter, Saturn, Uranus, and Neptune are composed almost entirely of gases, but the pressure deep beneath their cloudy exteriors causes the gaseous mixture to become extremely dense. Their atmospheres grade gradually into regions where the gases exist as liquids.

The chemistry of Jupiter's and Saturn's atmospheres is very similar, and abundances of hydrogen, carbon, and nitrogen are, in bulk, close to those of the Sun itself. These atmospheres are primordial. Both planets display a striking pattern of many alternating cloud bands lying parallel to their equators. Colors are primarily attributable to the photochemical reactions that occur at and near the cloud tops. Since Saturn lies about twice as far from the Sun as Jupiter, it receives significantly less solar radiation and ultraviolet energy. Photochemical reactions are not as strong, or as varied, as they are on Jupiter. The highest cloud level in Jupiter is a reddish color believed to be explained by the presence of phosphorus, which may be photochemically dissociated from phosphine gas known to be in Jupiter's upper atmosphere. The second level of clouds is white, the third layer is brown (most likely caused by sulfur), and the fourth layer is blue. Ammonia and ammonium hydrosulfide are major cloud-forming molecules present in Jupiter and Saturn. A better understanding of the relationship of the observed colors to the chemistry of the cloud layers must await the descent of instrumented probes into the atmospheres of the gas giants. From late 1995 to 2003, Galileo probe gathered data on Jupiter's upper atmosphere; its results did not match the expectations of the contemporary model for Jupiter's atmosphere.

Studies of features moving in the cloud bands show that the bands are driven by alternating patterns of east and west winds. Explaining this wind pattern presents a difficult challenge to planetologists. One leading theory holds that the fluid interior of each planet is organized as if it were a series of nested cylinders rotating around axes coincident with the planet's axis of rotation. If this is the case, the contra-rotating cloud bands are the extremities of the nested structures, visible to Earth at the point where they are truncated by the spherical limits of the planet itself. Laboratory studies show that such structures can develop in rapidly rotating fluids and that alternating "cylinders" rotate in opposite directions under certain conditions. Jupiter's Great Red Spot, Saturn's white spot, and numerous lesser cyclonic systems visible on both planets lend support to this theory of the origin and motion of the cloud bands. These features do not participate in the rapid westward or eastward motions of the cloud bands adjacent to them, but circle the planets more slowly, rolling like ball bearings between surfaces rotating in opposite directions. Moreover, if the cloud bands' motions are driven by fluid dynamics extending deep into the planet, the systems must have enormous inertia, explaining the observed persistence of the cloudy bands with very little variation over decades.

Although Uranus and Neptune are Jovian in nature and possess primordial atmospheres rich in hydrogen (84 percent) and helium (14 percent), their characteristics differ somewhat from those of the solar system's two largest planets. Among the suggested explanations for these differences are that Uranus and Neptune receive substantially less heat from the Sun, and both are physically much smaller than Jupiter and Saturn.

In Uranus's case, the planet's radical tilt on its axis and the resulting pattern of solar heating must be considered. Uranus displays a nearly featureless pale blue-green exterior in which there is almost no hint of the cloud bands typical of Jovian planets. This is caused by a high haze of methane and its photochemical byproducts. Spacecraft imaging has revealed that cloud bands are present beneath this haze, clearly indicating that the currents that drive the cloud bands must be caused by the planet's fast rotation rather than by solar insolation. Uranus was examined during an extended period of time when its northern hemisphere was exposed to continual sunlight and its southern hemisphere to continual night, yet the wind flow seemed unaffected by the unequal heating of the two. Nevertheless, it appears that Uranian cloud bands all rotate in the same direction, and there is no evidence of any large cyclonic structures being spun between contra-rotating bands.

Neptune's atmospheric features are both more distinct than Uranus's and more transient than those of any of the other gas giants. Cloud bands are visible even in the best terrestrial telescope images, and Voyager 2 discovered the Great Dark Spot in the planet's southern hemisphere, as well as several other smaller dark spots and bright features. These features tend to fade in and out, brighten and darken, and even wander across the latitudinal cloud bands, like cars changing lanes on a highway. The Great Dark Spot, evidently very similar to the big cyclonic features on Jupiter and Saturn, was observed to change size and shape as it rolled along between two contra-rotating cloud bands. Because of the limited amount of sunlight available to produce photochemical reactions at this distance from the Sun, the haze of methane and hydrocarbons is thinner at Neptune than at Uranus. Below the haze are clouds composed primarily of methane ice, and below that are layers of hydrogen sulfide and ammonia clouds.

The existence of an atmosphere around Pluto is somewhat speculative. If it truly exists as something more than just a tenuous feature, it must be very thin, since the planet is so cold. At 48 kelvins almost all gases except hydrogen and helium would form liquids or solids and settle on the surface of the planet as ice or snow. Also the planet is too small to retain hydrogen or helium. It is possible that a small amount of gaseous methane exists close to the surface. There is some evidence that the amount increases and decreases as the planet varies its distance from the Sun by 13 percent during the course of its orbit.

The only satellite in the solar system to possess a significant atmosphere is Titan, a satellite of Saturn whose diameter is about the same as that of Mercury. Data from the Voyager and Cassini missions confirmed a nitrogen-rich atmosphere as much as 200 kilometers deep, in which photochemical and electrochemical reactions have produced a high-altitude haze of complex hydrocarbons that hide all features of the surface and lower atmosphere. Of particular interest is evidence that polymer chains of organic molecules form in this haze and then sink to the surface of the planet. This scenario closely resembles the process proposed to explain the origin of life on Earth. However, the temperature of Titan, 94 kelvins, precludes liquid water, which is also believed necessary for life. Besides, Titan, some other satellites, especially those that are geologically active (and thus have conditions for outgassing), have tenuous atmospheres.

METHODS OF STUDY

Understanding planetary atmospheres beyond Earth has been facilitated greatly by data returned by such spacecraft as Mariner 10, a host of Venus and Mars probes, Pioneer 10 and 11, the two Voyagers, and the Galileo and Cassini orbiters. Prior to the epic journeys of these robotic laboratories, knowledge of the gaseous envelopes around the planets was severely limited by the difficulties of extracting meaningful data from faint images seen in telescopes.

Techniques for determining the chemical constituents of an atmosphere from afar are not essentially different whether the measurements are made through a telescope on Earth or from aboard a spacecraft coming within tens of thousands of kilometers of the planet. In both cases, the viewing telescope must be equipped with a spectroscope, an instrument that breaks the reflected light of the planet up into its component colors. It thereby reveals the "fingerprints" of chemicals present in the source of the light. A spacecraft in close proximity to a planet, however, has more reflected light with which to work and its target appears big enough that the instrument can be very discriminating about what it samples.

Coupled with the accuracy and specificity of the chemical assays provided by planetary space probes are the results of radio science experiments. These can probe the temperature and density of the deeper atmosphere. A wide variety of other instruments can image and sample the planet in many other wavelength regions of electromagnetic energy. In the cases of Venus and Mars, it has also been possible to send instrumented probes through the atmosphere, gathering data as they prepared to land on the surface.

Efforts to explore the solar system fully with sophisticated robotic spacecraft reached a peak during the interval between 1976 and 1989. Some planetologists have termed this the "golden age" because it produced an unprecedented harvest of information about the planets and converted a field of astronomy that previously had relied heavily on guesswork into a sophisticated branch of science, rich in data and theories with relevance to Earth. The revolution in planetary astronomy led to a new field of specialization called comparative planetology, which sought to learn more about the features (such as the atmospheres) of similar bodies by comparison to one another. A second age of robust planetary exploration began with advanced spacecraft in orbit about the planets such as Magellan and Venus Express (Venus), Galileo (Jupiter), and Cassini (Saturn). This period also included a sustained series of landers and orbiters sent to Mars on a fairly regular basis. Galileo and Cassini shed light on the atmospheric dynamics of the gas giants Jupiter and Saturn, respectively. They were follow-ons to the Voyagers, which provided tantalizing glimpses of these two planets as they flew past Jupiter and Saturn.

CONTEXT

Accompanying the great increase in data about the nature of the solar system has been a quantum improvement in the understanding of how each of the planets formed. This knowledge rests heavily on information concerning the planetary atmospheres themselves, particularly as the planets with primordial atmospheres retain vestiges of the chemical makeup of the solar nebula from which the Sun and all its planets were born.

Some fundamental questions regarding the atmospheres of the other planets remain unanswered. This is particularly true of the Jovian planets, since by 2008 only one entry probe, detached from the Galileo orbiter, had entered Jupiter's atmosphere. The depth to which the visible patterns extend is not known in general, nor is it clear whether and to what depth the atmospheres are layered below the cloud tops. Whether heat from the Sun or heat generated in the interior of the planet has the primary role in driving the circulation of the atmosphere has also not yet been determined. The next generation of planetary probes will seek to answer these questions.

Richard S. Knapp

FURTHER READING

Beatty, J. Kelly. "Getting to Know Neptune." *Sky and Telescope* 79 (February, 1990): 146-155. A synopsis of the findings from Voyager 2's encounter with Neptune in August, 1989. Several paragraphs discuss the atmosphere itself, and color diagrams enrich the text. The journal is intended for amateur astronomers and others with some general science background, but Beatty's expertise in science writing bridges the gap to the general reader.

Burgess, Eric. *Uranus and Neptune: The Distant Giants*. New York: Columbia University Press, 1988. Written after Voyager 2's Uranus encounter but before the spacecraft reached Neptune, this book's value is in its comprehensive discussion. Burgess presents both the findings of the spacecraft's instrumentation and the interpretation of these findings by scientists. Advanced reading.

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001. A challenging and thorough text for students of planetary geology, and an excellent reference for the most serious reader with a strong science background. Covers extrasolar planets and provides an in-depth explanation of the solar system's formation and evolution.

Esposito, Larry W., Ellen R. Stofan, and Thomas E. Cravens, eds. *Exploring Venus as a Terrestrial Planet*. New York: American Geophysical Union, 2007. This is a collection of articles covering all major areas of planetary research on Venus, including its atmosphere. Technical.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college astronomy text, somewhat more advanced than many introductory texts, with a wealth of detail and excellent diagrams. Chapters 6 through 16 describe the solar system. Comes with a CD-ROM.

Grinspoon, David Harry. *Venus Revealed: A New Look Below the Clouds of Our Mysterious Twin Planet*. New York: Basic Books, 1998. A thorough examination of the geology of Venus. Incorporates Magellan mapping and other data. Explains the Venusian greenhouse effect. A must for the planetary science enthusiast who wishes to read an integrated approach to science and history. Includes speculation about Venus's past.

Harland, David M. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002. Another book in Springer's Space Exploration Series, this is a technical description of the Cassini program, its science goals, and the instruments used to accomplish those goals. Written before Cassini arrived at Saturn. Provides a historical review of pre-Cassini knowledge of the Saturn system.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science. The chapter on Jupiter covers all aspects of the Jovian system and spacecraft exploration of it. Other chapters cover atmospheres on the other planets.

Harvey, Brian. *Russian Planetary Exploration: History, Development, Legacy, and Prospects*. New York: Springer, 2007. Early Russian space programs attempted a large number of Moon, Venus, and Mars investigations. Many were successful, many not. These robust programs are often overlooked. This is their story in one illuminating book about the engineering, development, flight operations, and science returns.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their atmospheres, as well as their satellites, rings, and magnetic fields. Filled with figures and photographs. For serious readers.

Leverington, David. *Babylon to Voyager and Beyond: A History of Planetary Astronomy*. New York: Cambridge University Press, 2003. A historical approach to planetary science. Heavily illustrated; concludes with a summary of spacecraft discoveries. Suitable for both general readers and the astronomy community.

Lorenz, Ralph, and Jacqueline Mitton. *Lifting Titan's Veil: Exploring the Giant Moon of Saturn*. Cambridge, England: Cambridge University Press, 2002. An in-depth examination of all that was known about Titan prior to the Cassini-Huygens mission, written by an engineer who worked for the European Space Agency and an astrophysicist who was press officer for the Royal Astronomical Society. Describes the Cassini-Huygens mission but not its outcome, as the book was published before the spacecraft arrived at Saturn.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar-system astronomy suitable for an introductory college course and nonspecialists as well. Filled with supplemental learning aids and solved student exercises. A Web site is available for educator support.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. Planetary atmospheres are treated as important physical features of the various members of the Sun's family. They are discussed individually in the context of what is known about each planet's characteristics and with regard to theories about their evolution and the evolution of the entire solar system. Geared for the undergraduate college student.

See also: Earth's Atmosphere; Jupiter's Atmosphere; Mars's Atmosphere; Neptune's Atmosphere; Saturn's Atmosphere; Uranus's Atmosphere; Venus's Atmosphere.

Planetary Classifications

Category: Planets and Planetology

Since ancient times, no formal scientific definition of the word "planet" existed. In 2006, professional astronomers for the first time developed a formal definition of a planet. The new definition removed Pluto, previously labeled a planet for seven decades, from the list of planets.

OVERVIEW

The word "planet" comes from the ancient Greek for "wanderer." From prehistoric times, people have looked up and seen stars. Five of those stars seemed to wander from constellation to constellation. The Sun and Moon also moved among the heavens. Thus seven bodies were historically classified as planets. They were the Sun, the Moon, Mars, Mercury, Jupiter, Venus, and Saturn. The seven days of the week became associated with these planets. Other objects were seen to move in the sky, such as comets and meteors, but they did not move in predictable paths, and astronomers even debated whether they were celestial or atmospheric phenomena.

Nicolaus Copernicus's heliocentric model, though, made the Sun the center of the solar system, with Earth, Mercury, Venus, Mars, Jupiter, and Saturn orbiting the Sun. The Moon orbits Earth. Proponents of the heliocentric model began listing planets as the bodies orbiting the Sun. Thus, six planets were recognized orbiting the Sun. When Galileo Galilei discovered four satellites of Jupiter, they, too, were at first listed as planets. Soon, however, it was decided that bodies orbiting other planets are not planets themselves, so the bodies that Galileo found were by acclamation declared to be moons, or satellites, of Jupiter. Eventually, astronomers discovered Uranus and Neptune orbiting the Sun, bringing the list of planets to eight.

In the late eighteenth century, astronomers Johann Titius and Johann Bode suggested that the gap between Mars and Jupiter was disproportionately large and that an unseen planet may lurk between the two. Then, on January 1, 1801, Giuseppe Piazzi found an object orbiting the Sun in this gap between Mars and Jupiter. The body was named Ceres, and it was declared to be a planet. In the next few years, though, astronomers also found Pallas, Vesta, and Juno, all orbiting between Mars and Jupiter. Each of these bodies was classified as a planet. As discoveries continued to mount, however, astronomers realized that it had become unwieldy to consider all of these bodies planets. Since the bodies between Mars and Jupiter were all so small as to not even show a disk in the best of telescopes, they were deemed too small to be considered planets and thus were unofficially declared to be asteroids. For most of the twentieth century, these bodies were called minor planets by professional astronomers.

Astronomers also came to recognize comets as bodies orbiting the Sun. By the twentieth century, comets were known to be icy bodies that originate in the outer solar system. When they are disturbed into highly elliptical orbits that come too close to the Sun, the ices and frozen gases that make up the comet sublimate and boil off into space along with dust to form the tail of the comet. A more modern understanding comets is that they may be similar to asteroids, only with a far larger percentage of ice and frozen gas.

In 1930, Clyde Tombaugh discovered Pluto. The planet Neptune had been found by observations of a perturbation in the orbit of Uranus by Neptune's gravity. Early in the twentieth century, several astronomers mistakenly thought that Neptune's orbit was also perturbed by another planet even farther from the Sun. Tombaugh was one of many astronomers searching for that planet when he discovered Pluto. Therefore, when Pluto was found, there was a predisposition to believe that it was quite large. Furthermore, at that time astronomers knew of only two types of planetary bodies: small rocky worlds, and gas giants shrouded in clouds. Pluto was clearly not a gas giant. If it were rocky, though, then it would have to be fairly large in order to appear as bright as it does, since rocky worlds tend to reflect only a small percentage of sunlight that hits them. Within a few years, however, Pluto was found to be icy in nature. Ice is far more reflective than rock, so an icy Pluto could be much smaller than a rocky Pluto.

By the 1970's, astronomers realized that Pluto is tiny. It has since been found to be only about one-fifth the diameter of Earth and about 0.2 percent the mass of Earth. Pluto is even smaller than Earth's moon. That makes Pluto much smaller than any other planet. Astronomers began debating as early as the 1980's whether Pluto should really remain on the list of planets, given that it is so small. Furthermore, Pluto's composition is more like that of comets, which also originate in the outer solar system, than that of the other planets.

However, serious proposals were put forth to remove Pluto from the list of planets only after other icy bodies, very much like Pluto only smaller, were discovered beyond Neptune. Then, in the early years of the twenty-first century, objects were found that were only slightly smaller than Pluto, making the classification of Pluto as a planet, while not considering objects only slightly smaller to be planets, problematic. Then object 2003 UB313 was discovered. This object, eventually named Eris, was somewhat larger than Pluto. Clearly, if Pluto were a planet, this object should also be a planet. The discovery of Eris, then, simply added to the arguments about what constituted a planet.

In 2005, therefore, the International Astronomical Union (IAU) established a committee to decide on a definition of what constitutes a planet. The committee made a recommendation that any body orbiting the Sun whose gravity is sufficient to form the body into an essentially spherical shape should be declared a planet. Such a definition retained Pluto's status as a planet, added Eris to the list of planets, and returned Ceres to planetary status. However, this definition also made about a dozen other bodies possible planets; their planetary status would be unknown until astronomers were able to determine their shapes.

The committee's proposal was rejected by the IAU. In 2006, after some deliberation, the IAU general assembly adopted a different definition of planet. That definition states that a "planet" is a body orbiting the Sun large enough to form itself essentially into a sphere and sufficiently large for its gravity to clear the neighborhood of its orbit. A new category, called a "dwarf planet," was defined as a body large enough to form itself into a spherical shape but not sufficiently massive to clear its orbital neighborhood. All other objects were to be designated as "small solar system bodies." Since

The IAU's 2006 Classification System

On August 24, 2006, the International Astronomical Union redefined bodies in the solar system, including planets, as follows:

The IAU . . . resolves that planets and other bodies, except satellites, in our Solar System be defined into three distinct categories in the following way:

(1) A planet[1] is a celestial body that (a) is in orbit around the Sun, (b) has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape, and (c) has cleared the neighborhood around its orbit.

(2) A "dwarf planet" is a celestial body that (a) is in orbit around the Sun, (b) has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape,[2] (c) has not cleared the neighborhood around its orbit, and (d) is not a satellite.

(3) All other objects,[3] except satellites, orbiting the Sun shall be referred to collectively as "Small Solar System Bodies."

The IAU further resolves: Pluto is a "dwarf planet" by the above definition and is recognized as the prototype of a new category of Trans-Neptunian Objects.[4]

Notes:

1. The eight planets are: Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, and Neptune.
2. An IAU process will be established to assign borderline objects into either dwarf planet [or] other categories.
3. These currently include most of the Solar System asteroids, most Trans-Neptunian Objects (TNOs), comets, and other small bodies.
4. An IAU process will be established to select a name for this category.

Pluto is only one of many objects of similar size and has not cleared its orbit, it was reassigned as a dwarf planet. Likewise, Ceres and Eris are dwarf planets. The term "asteroid," never an official term, was not mentioned, nor was the term "comet." Rather, all comets, asteroids, and trans-Neptunian objects not large enough for their gravity to pull themselves into roughly spherical shapes are lumped together in the category of small solar system bodies.

KNOWLEDGE GAINED

Scientists classify objects in order to aid their study of them. Astronomers are no different from other scientists in that respect. Classifying celestial bodies helps to organize knowledge about those bodies. Knowing that a newly discovered body is a comet or an asteroid gives a mental picture of its physical characteristics. However, the classification of objects as planets, comets, asteroids, and other entities was never officially defined before 2006. Astronomers had simply agreed informally on their definitions, and as a result some bodies did not obviously fall into one category or another. With the discovery of more and more objects that fell between classifications—such as icy asteroids and not-so-icy comets—distinctions became more difficult. Eventually objects such as Pluto, Eris, and similar sized icy bodies forced a formal definition. However, both the original proposed definition and the one finally approved by the IAU have been criticized as flawed, and with future discoveries and increasing data and insights, the issue may be eventually decided again.

There are multiple ways to classify objects in the solar system. Objects can be classified by size, by method of formation, and by physical composition. If size is the primary criterion for defining a planet, the eight bodies now defined as planets do indeed meet that criterion, since they are substantially larger than the other solar system bodies. If method of formation is the most important factor, then Jupiter and Saturn would be seen differently, since they formed in a manner different from thaat of the other bodies. Certainly, the inner, rocky planets, formed from components different from those of the outer planets; most astronomers consider the inner planets to be formed from smaller bodies such as the asteroids, which are likely leftovers from the formation of the solar system. The outer planets Uranus and Neptune are believed to have been formed from objects more like Pluto and the other trans-Neptunian bodies, making those objects in the outer solar system analogous to the asteroids in the inner solar system.

A major criticism of the 2006 IAU definition of planet is that it does not define what is meant by a planet "clearing the neighborhood of its orbit." For example, Jupiter shares its orbit with a number of asteroids, but those asteroids' orbits are dominated and determined by Jupiter's gravity. Asteroids in elliptical orbits cross the orbits of Earth, Mars, and Venus, yet those bodies are considered planets. The asteroids that cross those planetary orbits are unstable and cannot continue to cross those orbits forever. However, the orbits of the inner planets have not been completely cleared yet. Furthermore, the degree to which a body must be deformed toward a spherical shape to become a dwarf planet is not well defined. There are many bodies that are somewhat, but not quite, spherical. It is unclear if they are to be considered dwarf planets or small solar system bodies.

CONTEXT

Classification of bodies in the solar system is important to textbook authors and students, who often have as their first introduction to the solar system memorization of the planets and the types of objects in the solar system. There are different naming schemes for the different types of bodies, too, so knowing the classification for an object determines what sort of name is permitted. Planets are named one way, comets another way, and asteroids in yet a third way.

However, the particular classification of a body such as Pluto does not in any way change how professional astronomers view it. For many years, even before its reclassification as a "dwarf planet," astronomers regarded Pluto as more like a comet or asteroid than like a planet. With the discovery that other trans-Neptunian objects were similar in size and composition to Pluto, astronomers began thinking of Pluto as simply one of the largest of those bodies. How it is classified in textbooks does not change how they regard it.

Making a rational definition of a planet is important to scientists in properly grouping bodies together that have similar characteristics in order to facilitate study of those objects. When the scientific understanding of a body changes, then it sometimes needs to be reclassified. This reclassification is based on scientific principles, not public opinion. In the aftermath of the 2006 IAU definition, an attempt to mollify public affection for Pluto as the "ninth planet" led to an another proposal for a definition of "planet" that did little to clarify the issue in the end. The final definition was crafted hastily and without sufficient forethought, making it, too, an incomplete definition. Therefore, the definitions of planets and other bodies may well be revisited in the near future.

Raymond D. Benge, Jr.

FURTHER READING

Corfield, Richard. *Lives of the Planets: A Natural History of the Solar System*. New York: Basic Books, 2007. A survey of the solar system focusing primarily on the knowledge gained through robotic spacecraft exploration of the solar system. Several color photographs are included.

Soter, Steven. "What Is a Planet?" *Scientific American* 296, no. 1 (January, 2007). This article does a very good job of explaining several possible definitions of what constitutes a planet. It includes a brief historical perspective.

Stern, Alan, and Jacqueline Mitton. *Pluto and Charon*. New York: Wiley, 1999. An easy-to-follow history of the study of Pluto. Discusses what is known about Pluto and its satellite Charon without being too technical. Also covers the other icy, Pluto-like bodies beyond Neptune.

Tyson, Neil deGrasse. "A Death in the Solar System." *Discover* 27, no. 11 (November, 2006). A clear and simple explanation of the IAU's definition of a planet. Includes several photographs and tables.

Weintraub, David A. *Is Pluto a Planet? A Historical Journey Through the Solar System*. Princeton, N.J.: Princeton University Press, 2007. As indicated in the subtitle, most of the is about the history of astronomical discoveries in the solar system, with only the last chapters directly addressing an answer to the question of Pluto's status as a planet.

Woolfson, Michael. *The Formation of the Solar System: Theories Old and New*. London: Imperial College Press, 2007. A concise but thorough survey of the theories of planetary formation. Though equations are included, mathematical prowess is not required to understand the discussion.

See also: Planetary Formation; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Pluto and Charon; Terrestrial Planets.

Planetary Formation

Category: Planets and Planetology

The processes that influence planetary formation are extremely complicated but reasonably well understood. In the densest portion of a contracting solar nebula, a star forms, and, assuming that conditions are suitable, a family of planets can also form around that star. Planet formation may be a very common occurrence throughout the galaxy.

OVERVIEW

"Planet formation" is the term used to describe how various planetary bodies formed from residual material surrounding a forming star. The two most discussed theories suggest either a catastrophic or an evolutionary origin for the planets. In either case, the planets would have formed in a manner similar to the way the star formed, but on a much smaller scale and with considerably less mass.

One theory suggests that the formation of a solar system begins with the explosion of a massive nearby star, called a supernova. A supernova begins with a slowly contracting cloud of gas and dust that is capable of forming many individual stars, called a cocoon nebula. Where the concentrations of matter are densest, giant stars will form. If they are massive enough, they will progress rapidly through their evolutionary stages and explode with gigantic force. Energy released may rival that of a galaxy for a short period of time. Compressional shock waves generated by the explosion, along with the excess matter thrown off, can initiate formation of smaller nearby stars such as the Sun. Evidence for such an event is suggested by certain isotopic ratios for various elements present in meteorite compositions and even the Sun itself.

The majority of mass in a solar nebula will go into the formation of individual stars. Some estimates indicate that a star will account for greater than 99 percent of the total mass. The remaining 1 percent will represent matter available for planetary formation. Increasing pressures and temperatures at the gravitational center of a solar nebula will eventually form a protostar. This is the initial stage that precedes the beginning of nuclear processes in a star. The protostar is more representative of a hot glowing ball of gas than a true star. It is believed that planets will commence their earliest formation processes at this time as well.

Planet formation is characterized by three distinct processes: gravitational collapse, condensation, and accretion. The initial process that took place in the solar nebula was gravitational collapse, the process in which atom attracts atom to form molecules of considerably greater mass. This greater mass increases gravitational pull and thereby attracts more atoms. As these localized regions became denser through the infall of matter, temperatures began to increase. Gases and solid dust grains that were originally cold now begin to heat up and glow as more and more matter becomes concentrated in a rapidly shrinking volume. Scientists believe that eventually all matter will be converted into a hot gaseous state.

After the majority of gas has been incorporated into the star's formation, residual hot gases begin to cool and re-form as solids during the condensation process. Every element has a specific condensation temperature at which a phase change takes place. Metallic elements such as aluminum and magnesium condense from a gaseous state at very high temperatures, while gases such as methane or water vapor condense at low temperatures. Where concentrations of matter were greatest near Earth's proto-Sun, for example, only heavy elements condensed. As matter thinned out, temperatures dropped progressively and permitted lower-temperature elements to condense. Based on existing temperatures, distance from the protostar controlled, in part, the abundance ratios of the various elements found in those regions. This is one reason that the inner planets are high in iron content relative to the outer gaseous planets.

Material in the residual gas rings that surrounded the proto-Sun was pulled to the plane of the ecliptic by the force of gravity. This material tended to clump together as it condensed out of the cooling gases. Some scientists believe that in our solar system this condensation phase was rapid enough to trap some of the lighter gases within the forming clumps of heavier elements. This may account for the large amounts of volatiles present during the final stages of planetary formation.

As condensation progressed, only microscopic grains were in evidence. Later, these combined to form larger and larger particles. Growth continued through sand-sized to pebble-sized and then to boulder-sized objects. This describes the accretion process, where collision produced a larger mass rather than fragmentation. Accretion did not stop with boulders, but continued, eventually forming planetesimal-sized objects much like present-day asteroids. They had a size range from 30 meters up to about 1,000 kilometers in diameter. Finally, planetesimals collected one another and assumed the planetary dimensions familiar today.

Present-day planets are not uniform in their chemical compositions. The inner, so-called terrestrial planets (Mercury, Venus, Mars, and Earth) are rich in heavy metals, while the outer planets (Jupiter, Saturn, Uranus, and Neptune) are gaseous—hence, the so-called "gas giants." One explanation for this may be that the solar nebula was not uniformly hot, because of nonhomogeneous distribution of material based on distance from the proto-Sun. Only the area from the Sun outward to the asteroid belt (a region of small rocky and metallic bodies between the orbits of Mars and Jupiter) was sufficiently hot to experience heating of the original nebular material and subsequent condensation. Material that was present at the distance of Jupiter and beyond remained in its unaltered state until it became concentrated in the protoplanetary stage of formation.

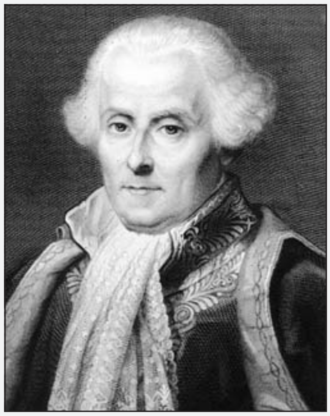
An additional factor to be considered in planetary formation is the T-Tauri stage in the development of a main sequence star like the Sun. This is a hyperactive stage in a young star's life, as fusion reactions begin to dominate. As a result, large quantities of matter are ejected at a very rapid rate. It has been suggested that the early Sun may have lost as much as one-half its original mass in as short a period as 1 million years. This would be evident as an intense solar wind many times greater than it is today.

Distance would once again be an influencing factor in planet formation. As the solar wind came in contact with the inner protoplanets, it would have stripped away the lighter elements, thereby concentrating the heavier ones. Yet, even with its great intensity during the T-Tauri stage, the solar wind would have lost most of its effectiveness at the distance of the present asteroid belt. It may have been virtually ineffective by the vicinity of Jupiter, thereby leaving the outer planets with most of their original mass. In fact, the Jovian planets may have increased their mass as gravity collected additional matter flowing past from the inner solar system.

Pierre-Simon Laplace: Solving the Unsolvable

Pierre-Simon Laplace was born on March 23, 1749, to a well-off farming family in Normandy who could provide their children with educational opportunities. He studied mathematics at the University of Caen, taught at his former priory school, and then met the famous mathematician Jean le Rond d'Alembert, who in 1770 recommended him to the École Royale Militaire in Paris.

Turning to astronomy, in 1773 Laplace solved a problem in celestial mechanics that had puzzled many scholars, including the famous Leonhard Euler and Joseph-Louis Lagrange: Apparent variations in the speeds at which the planets revolved around the Sun seemed to have no reasonable explanation and indicated a worrisome instability in the solar system. Sir Isaac Newton himself had concluded that the solar system required intermittent divine intervention to keep it going. Laplace demonstrated that this planetary instability was only apparent; in fact, the variation in the speed of the planets was a periodic phenomenon that could be predicted. This work opened the door to the Académie des Sciences, and he would become its president in 1812.



(AIP Niels Bohr Library)

Laplace's scientific reputation placed him in an intellectual, social, and political position that enabled him to have a profound influence on the changes that occurred in France during the revolution of 1789. In 1790, Laplace was instrumental in developing the metric system, and in 1795 he was a cofounder and first director of the Bureau des Longitudes. He also served as the director of the Paris Observatory. In 1806, Napoleon named Laplace a *comte d'empire* and gave him the position of minister of the interior in his government. Much more a scientist than an administrator, Laplace kept his ministerial position for only six weeks before gracefully withdrawing. In 1807 Laplace, along with the well-known chemist Claude Louis Berthollet, organized a group of famous scientists and young researchers called La Société d'Arcueil to encourage promising young graduates of the Polytechnique to continue their research.

Laplace was elected to the French Academy in 1816, and in 1817 Louis XVIII named him a marquis. When he died at the age of seventy-eight, Laplace not only had revolutionized astronomy but also had profoundly transformed the fields of probability, weights and measures, and mathematics. In a eulogy presented at the French Academy on November 13, 1827, Laplace's successor, Pierre-Paul Royer-Collard, said:

> Laplace was born to perfect everything, to deepen everything, to push back the limits, and to solve all of the things people believe unsolvable. He would have completed the study of astronomy if this science could be completed.

In the end, the planets as they are known today were formed. The processes at work were very complex, involving many variables. Collectively, the aforementioned theories are known as the nebular hypothesis. It offers reasonably logical explanations for a very difficult problem.

METHODS OF STUDY

Experimentally, several aspects of the nebular hypothesis can be examined in the laboratory through high-temperature and pressure studies on various materials believed to have been present in the nebula. Nevertheless, in the final analysis, no experiment can reproduce an actual solar system or a planet. It is possible, however, to compile all available data into a computer model and develop a graphic representation of what may have taken place.

Scientists do not have to rely totally on computer modeling for answers; there are actual clues. Evidence for the condensation and accretion phases of planet formation can be found in the chemistry and mineralogy of stone meteorites. Chemical compositions of minerals found in meteoritic objects called chondrules bear witness to an early high-temperature condensation history. The physical condition of these chondrules—as they are found in the meteorite—also exhibits evidence of a very violent phase during which collisions were commonplace. These normally spherically shaped objects can be found as fragments or in a partially melted and deformed state. They were incorporated later into a re-forming object that eventually would reach planetesimal proportions. Although scientists do not totally agree about the mechanisms for the origin of chondrules, they are certain of their primitive nature.

The bulk chemistry of meteorites also provides additional information about planetary formation. One type of meteorite, called a type one carbonaceous chondrite, contains the most primordial chemistry of all materials in the solar system. Its chemical composition most closely resembles the "condensable" part of the Sun's chemical makeup. Also present are low-temperature, water-bearing minerals; carbonate and sulfate minerals; and 8 percent to 22 percent water. High-temperature minerals are rare, and chondrules are notably absent. This strongly suggests that this type of meteorite may be very close in chemical composition to the original unaltered material of the solar system.

The unique primordial nature of meteorites was destroyed as they were incorporated into the growing planetesimals. As the planetesimals accreted further into actual planet-sized masses, evolutionary processes took over, erasing all evidence of their protoplanetary stage. For the inner planets, intense bombardment from infalling planetesimals turned their surfaces into a crater-dominated world. Those planets that had sufficient mass would heat up internally, with volcanism, plate tectonics, and erosion eventually reshaping their surfaces. In contrast, the outer planets never got beyond their protoplanet stage. Their great masses kept them hot, and they never cooled off.

The first extrasolar planet was discovered in 1995 using a technique that involves measuring the gravitationally induced wobble in a star produced by a planet orbiting it. Ironically, the first planet found outside the solar system did not fit the typical model envisioned for a planetary system. It was a very large gas giant in an extremely tight orbit about a pulsar. Pulsars had not been considered likely candidates for solar systems with planets. A decade after that initial exoplanet discovery, more than one hundred other planets had been detected through a small number of indirect means. Most extrasolar planets were hot Jupiters, but some were found having masses down to a multiple of a Uranus-class mass. After the first decade of discoveries, the number of objects beyond the solar system that were accepted as planets rose dramatically. By 2008, there were more than three hundred recognized exoplanets. The tremendous variety of these planets and the nature of the solar systems in which they were found forced a serious reexamination of the theory of planetary system formation.

METHODS OF STUDY

The nebular hypothesis for planetary formation provides good explanations for many of the dynamical and chemical properties found in the solar system, but it is not all-inclusive. Some inconsistencies do exist, such as the axial tilt and the rotational direction of Uranus. The possible formation of this planet is more representative of a catastrophic event that may have literally knocked it on its side.

Comets are another example. Although much is known about their physical and chemical characteristics, little is known about how they fit into the nebular hypothesis. They apparently come from the outer fringe of the solar system, but their highly elliptical orbits are not consistent with protoplanetary theory. A likely explanation would suggest that a catastrophic event is somehow also responsible for the origins of comets.

One place scientists can look for possible confirmation of the nebular hypothesis is in the satellite systems of the gaseous planets. At their formation these protoplanetary bodies resembled small stars rather than planets. It is possible that they acted in a manner similar to a forming star with respect to their satellite families. One can envision a protoplanetary Jupiter gravitationally gathering in the largest amount of mass and perhaps radiating a small amount of energy just as the early proto-Sun did. Its mass would not permit it to progress further, but it may have remained hot long enough to influence the chemical compositions of its inner higher-density satellites (Io, density 3.5 and Europa, density 3.0). In contrast, the two larger outer satellites, Ganymede and Callisto, have densities of 1.9 and 1.8, respectively. This comparison between the satellites of Jupiter and the arrangement of planets by density is inescapable. The inner two Jovian satellites are composed of rock, ice, and small amounts of metal, while the outer two are dominated by ice and dust. Astronomers are investigating if they truly formed according to protoplanetary theory. Only future spacecraft exploration will tell.

The ring systems of the outer planets may well answer some of the questions concerning accretion and planetesimal formation. For example, are the rings representative of a fragment satellite, or one just coming together? What about orbital dynamics and the influence of shepherding satellites on the orbital motion of the particles in the rings? These are all valid questions whose answers will undoubtedly shed light on planetary origin.

As one moves outward through the solar system, the density comparison between satellites in a given system is not so apparent as it is in the Jovian system. These satellites all appear to be relatively uniform mixtures of various ices and dust. Even large Titan, with a density of 1.9, does not match up to Io and Europa. It is then apparent that temperatures in these far-removed regions of the solar nebula were merely too cold to permit any large concentration of heavier elements, regardless of the influence of the planet they orbit.

Perhaps the most striking evidence for accretion can be found in the physical characteristics of the small Uranian moon called Miranda. Apparent in the Voyager 2 photographs were two distinctly different surfaces. One was a densely cratered terrain typical of the low-density ice satellites of the outer planets. The other had no comparison. There were circular patterns of angular landforms surrounded by alternating bright and dark bands, along with fault scraps and ridges. These were clearly of a much younger age than the ancient cratered surface. Scientists are searching for a possible explanation for such a unique satellite. One theory suggests that Miranda may have experienced an early partial internal melting. This would have been followed by a low-velocity impact that broke it in several pieces. These large pieces remained in orbits close to one another and in a relatively short period of time reaccreted into a single large mass. The patchwork nature of Miranda's surface makes this a strong possibility and would thus provide astronomers with an example of an actual accretionary body.

Theory and the examples from the solar system can provide valuable evidence for planet formation, but scientists would like to see the actual event as it happens. Considering the vast number of young stars in the Milky Way, this should be possible, but distance prevents examination by earthbound telescopes. Methods of observation that make use of non-visible radiation, however, may provide the information necessary to confirm the nebular hypothesis.

In 1983, the Infrared Astronomical Satellite (IRAS) began gathering data on various sources of infrared energy in space. This included energy emitted from newly forming planetary systems. One of the first candidate sources was the young star Vega. IRAS showed it to be surrounded by a spherical cloud of cool dust, just as theory predicted for a forming planetary system. A second observation showed the star Beta Pictoris to have a dust cloud surrounding it in a typical flattened plane. Optical observations confirmed this disk to have a radius ten times larger than that of Pluto's orbit and confirmed that the dust was orbiting within 5° of that star's plane of the elliptic. Evidence such as this does not necessarily confirm the nebular hypothesis, but it strongly supports it.

The Hubble Space Telescope and Spitzer Space Telescope have since found many more protostellar objects and forming planetary systems. The indirect detection of more than three hundred extrasolar planets (a number that is continually increasing), the vast majority of which do not appear to have formed in a manner consistent with the model of our solar system condensing out of a planetary nebula, puts constraints on the acceptability of the nebular hypothesis, or at least its universality. "Hot Jupiters" close to a star, pulsars with planets, and a few other unusual star-planet relationships fall well outside the classic planetary nebula model.

Much anticipated is the first direct imaging of an Earth-class planet about another star. The Kepler Observatory, designed to detect planets while transiting their star, was expected to deploy in 2009; if successful, it would survey as many as 100,000 star systems. It is estimated that the chances for finding a terrestrial planet are on the order of 1 in 210; if that low probability proves to be correct, then Kepler might detect as many as 476 Earth-class planets. Such continuing indirect and direct observations of existing and forming planetary systems will surely require major modifications to the theory of planet formation.

In mid-June, 2008, European astronomers announced remarkable results found using the European Southern Observatory (ESO) La Silla Observatory's HARP spectrograph. Surrounding the star HD40307 were found three super-Earth-scale planets. HARP was capable of detecting planets between two and ten Earth masses—thus, objects on a planetary scale, between Earth and Uranus or Neptune in mass. This technique was sensitive to stellar wobbles of only a few meters per second, something necessary to determine the tiny tug on a star by an Earth-sized extrasolar planet at significant distances from its star. These researchers also announced that they expected to refine data on forty-five candidate Earth-scale exoplanets with masses between 1 and 30 Earth masses as they continued to observe with HARP. However, an extrasolar planetary system with three Earth-class planets was unique thus far in the history of exoplanet identification.

CONTEXT

Prior to the mid-seventeenth century, most theories for the origin of Earth and the solar system were based on either myth or religious interpretation. Later, science would tend to take either an evolutionary or a catastrophic approach. It was realized quickly that no comprehensive answer would be forthcoming, and the "best-fit" concept was developed to provide answers based on the best science available.

In 1644, French philosopher René Descartes presented the first scientific theory for the origin of the solar system. His model was based on an evolutionary process. Descartes envisioned space as filled with swirling gases. Large whirlpools developed that would form stars eventually, with smaller vortices forming planets and their satellites. In 1755, Immanuel Kant applied Sir Isaac Newton's laws of gravity to Descartes's model and concluded that the swirling gases would assume a disk shape. In 1796, the French mathematician Pierre-Simon Laplace proposed further that the disk would separate eventually into rings. The planets would condense from these rings. Collectively, the theories of Descartes, Kant, and Laplace became known as the nebular hypothesis. It offered reasonable explanations for the flattened appearance of the solar system, the nearly circular orbits of the planets, and their similar orbital directions around the Sun. Nevertheless, it left many questions unanswered and by the late nineteenth century had fallen from popularity.

In 1745, French naturalist Comte de Buffon proposed a catastrophic explanation that contrasted with the nebular hypothesis. His theory suggested that the solar system formed after a collision between the Sun and another massive object. The material thrown out by that collision eventually formed the planets. It was a relatively simple theory that provided answers to questions left unanswered by the nebular hypothesis. The Buffon theory was resurrected and modified by American geologist Thomas Chronder Chamberlin and American astronomer Forest Ray Moulton in 1900, and in 1917 it was revised by Sir James Hopwood Jeans. Jeans's revision had no collision but instead envisioned matter being pulled out of the Sun by the tidal forces of a passing star. This matter would then condense to form planets. Jeans's theory answered certain questions that the nebular hypothesis could not, but it was not a complete explanation.

Technological advances made during the twentieth and twenty-first centuries gave the nebular hypothesis new credibility. From the work of Carl von Weizsäcker, Gerard Peter Kuiper, and Hannes Alfvén, a revised model of the nebular hypothesis arose, envisioning a vast interstellar cloud of gas and dust that fragmented and contracted into smaller, dense regions from which stars would form. It is from within one of these smaller regions that the Sun and its planets formed.

In astronomical terms, the formation of the Sun and planets happened relatively quickly, perhaps in only 100 million years. It is suggested that zones of turbulent eddies once existed in the gas and dust cloud. Through a combination of processes of regional gravitational collapse, condensation, and subsequent accretion of solid particles, large protoplanetary bodies approximated the chemical compositions and were probably the precursors of the present-day planets.

Theories of planetary formation have given humankind a better understanding of the cosmic origins and the intricacies that led to Earth's existence in the solar system. In seeking answers to such questions, both mental and technological capabilities have been expanded, and this has overflowed into everyday lives. Experimentally, the biggest goal of the search for extrasolar planets is the detection and later the imaging of Earth-like planets.

Paul P. Sipiera

FURTHER READING

Arny, Thomas T. *Explorations: An Introduction to Astronomy*. 3d ed. New York: McGraw-Hill, 2003. A general astronomy text for the nonscientist. Includes an interactive CD-ROM and is updated with a Web site.

Brush, Stephen G. *Nebulous Earth: The Origin of the Solar System and the Core of the Earth from Laplace to Jeffreys*. Cambridge, England: Cambridge University Press, 1997. Describes how theories about the origin of the solar system have changed over time.

Consolmagno, Guy, and Martha Schaefer. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994. A text accessible to college-level science students and astronomers. Most subjects are presented using low-level mathematics, although integral calculus is used where required. Demonstrates how the area of planetary science progresses by questioning our previous understanding in light of new observations.

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001. A challenging and thorough text for students of planetary geology, serving as an excellent reference for the most serious reader with a strong science background. Covers extrasolar planets and provides an in-depth contemporary explanation of the formation and evolution of the solar system.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through space missions that had investigated all planets with the exception of Pluto by the publication date. Takes an astrophysical approach to place our solar system in a wider context as one member of similar systems throughout the universe.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college text on astronomy, somewhat more advanced than many introductory texts but with a wealth of detail and excellent diagrams. Chapters 6 through 16 describe the solar system. Comes with a CD-ROM.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science.

Hartmann, William K., and Chris Impey. *Astronomy: The Cosmic Journey*. New York: Brooks/Cole, 2001. A comprehensive survey of astronomy. The reader is taken on a journey through cosmic time from the big bang to the ultimate fate of the universe. Comes with a CD-ROM and other learning aids.

Short, Nicholas M. *Planetary Geology*. Englewood Cliffs, N.J.: Prentice-Hall, 1975. Although this text is somewhat dated, it still serves as an excellent overview for planetary studies. Chapter 4 discusses the origin of planets and presents a very good description for the nonscientist.

See also: Ceres; Dwarf Planets; Eris and Dysnomia; Extrasolar Planets: Detection Methods; Jovian Planets; Planetary Classifications; Planetary Satellites; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Red Dwarf Stars; Terrestrial Planets.

Planetary Interiors

Category: Planets and Planetology

Planetary interiors characteristically have a nested shell structure, with concentric layers of differing physical properties and chemical compositions. Often, planetary materials are compressed into unusual forms by the tremendous pressures deep within planets.

OVERVIEW

The interior of Earth is typical in many ways of the interiors of other terrestrial, or "rocky," planets and serves as a useful starting point for understanding those planetary interiors. Earth consists of three layers, differing sharply in composition and physical properties: a thin outermost layer called the crust, a thick intermediate layer called the mantle, and a dense central core.

Earth's crust consists of two quite different materials: granitic rocks averaging about 40 kilometers thick beneath the continents and basaltic rocks 5 to 10 kilometers thick beneath the ocean basins. Sea waters fill the ocean basins because the crust is thin there and its elevation low, but there is no direct connection between oceans and oceanic crust. The crust of Earth composes only about 0.3 percent of its mass and about 0.8 percent of its volume. Beneath Earth's crust is a dense mantle of iron and magnesium silicate rocks extending to a depth of about 2,885 kilometers. The mantle occupies about five-sixths of the volume of the Earth and accounts for two-thirds of its mass.

At the center of Earth is a dense core consisting of two parts. The outer core, with a radius of 3,486 kilometers, is slightly more than half the diameter of Earth and about the size of the planet Mars. This outer core is most likely molten iron and nickel with other elements present in smaller amounts. Fluid movements in the liquid outer core are believed to be the source of Earth's magnetic field. The inner core, which is solid but similar to the outer core in composition, has a radius of 1,216 kilometers. The core occupies about one-sixth of the volume of Earth and accounts for a third of its mass. The inner core accounts for only 0.7 percent of the volume of Earth and 1 percent of its mass.

The composition of Earth's interior, as well as that of other planets, is dictated by the chemistry of the solar system. If the solar system formed from a cloud of material similar to the present Sun in composition, any solid grains that condensed should have a composition similar to the Sun, minus those elements that would have formed gases. In the inner solar system (above the freezing point of water), most hydrogen, nitrogen, and carbon would have formed gases. Oxygen would largely be tied up in water vapor, but much would form oxides with various metallic elements or combine with silicon to form silicate minerals. Some sulfur would be available for iron sulfide minerals such as troilite (FeS). Noble gases (helium, neon, argon, krypton, and xenon) combine rarely with other elements and would be nearly absent from any solid grains. The most abundant elements of the inner planets should be oxygen, magnesium, silicon, sulfur, and iron. Direct chemical analysis of solar system bodies is still largely limited to Earth, the Moon, meteorites, and more recently Mars, but these analyses are consistent with theoretical predictions.

Meteorites are particularly valuable because they provide samples of the materials that condensed in the early solar system. A class of meteorites called chondrites closely matches the theoretically predicted composition of the planets; they are believed to be samples of very primitive solar system meteorites. Chondrites are usually the starting point for studies of the chemical composition of the inner planets. The presumed composition of Earth's mantle, mostly iron and magnesium silicates, agrees well with the theory that Earth has an overall chondritic composition. Much of the iron in Earth has settled into the central core, together with nickel and perhaps much of Earth's sulfur.

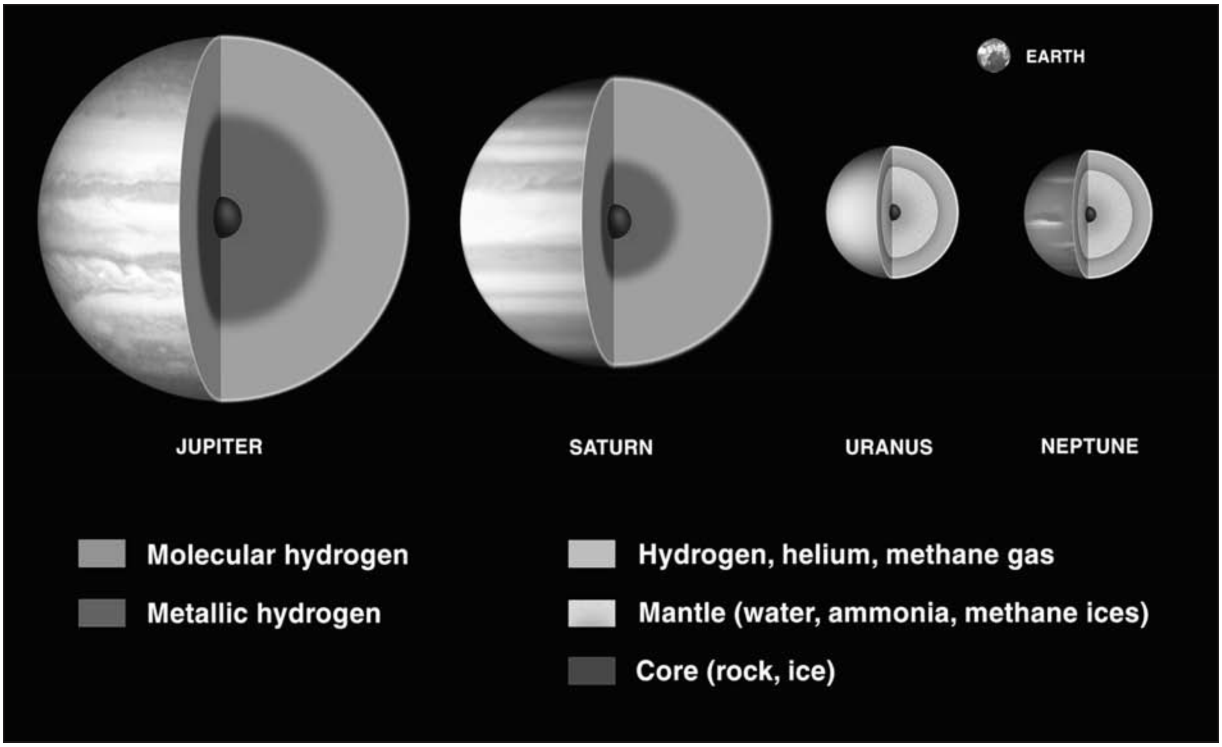
Pressure deep within the Earth is caused by the weight of overlying rocks. It is easy to calculate once the distribution of mass within Earth is known. The gravitational attraction within Earth is solely caused by the mass of Earth between the observer and the core; the gravitational effect of a shell surrounding an observer is zero. In a uniform Earth, gravitational attraction would decrease steadily toward the center, but because Earth's core is dense, gravitational attraction actually increases slightly with depth and is about 4 percent higher at the boundary of the core than at Earth's surface. Pressure at the base of Earth's continental crust is about ten thousand times greater than Earth's atmospheric pressure. Geologists commonly use a metric unit, the bar, to measure pressure. A bar is convenient because it is nearly equal to the pressure of Earth's atmosphere at sea level and thus is easily visualized. A bar is equal to 100,000 pascals. The pressure at the bottom of the mantle is about 1.4 million bars, and the pressure at Earth's center is about 3.6 million bars.

Temperatures deep inside the planet are not known as precisely as many other physical quantities. Earth becomes hotter at an average rate of 25 kelvins per kilometer near the surface, reaching about 1,250 kelvins at the base of the continental crust. Thereafter, temperatures increase much more slowly, reaching 3,000 to 5,000 kelvins in the core. Despite high temperatures in the mantle, it is still solid, because high pressure elevates the melting point of rocks. Portions of the mantle melt at depths of 50 to 100 kilometers only locally to generate magma for volcanic eruptions. Much of Earth's internal heat is caused by radiogenic heating, but some may be primordial remnants from the accretion of Earth or gravitational energy released when dense material sank to create the core.

Higher temperatures in Earth's deep interior create convection, the rise of hot, light material and the sinking of cooler, denser material. Although the rocks in the mantle are very rigid over short time spans, they can flow slowly over long timescales. It is still not known whether convection occurs throughout the mantle, whether there is one layer of convective motion or several, or precisely how the changes in crystal structure of mantle rocks at different depths might affect convection. Crustal plates consist of the crust, plus part of the mantle beneath. The material of the plates, collectively called the lithosphere, is about 100 kilometers thick.

Movement of crustal plates on Earth results in recycling of oceanic crust over a period of about 200 million years. New oceanic crust is formed as plates spread apart at the mid-ocean ridges, and old oceanic crust returns to Earth's interior at oceanic trenches, mostly around the Pacific rim. As oceanic crust is drawn back into Earth's interior, it melts partially, and molten material invades Earth's crust. This molten material is enriched in elements of large ionic radius and large electric charge that do not fit easily into the crystal structures of mantle rocks. During Earth's lifetime, the crust has become enriched in potassium, rubidium, silicon, uranium, thorium, and rare Earth elements. As a result the mantle has become depleted. Volatile materials such as water and carbon dioxide have also escaped to the surface.

Physical processes active in Earth's interior may be at work in other planets as well: internal heating, recrystallization of rocks and other materials to denser forms at great depths, gravitational separation of a dense metallic or metallic sulfide core, internal flow caused by convection, and segregation of particular materials



Interiors of gas giant planets. (Lunar and Planetary Institute)

into a crust. These processes are all driven, directly or indirectly, by internal heat. Small planetary bodies have less internal heat than large ones. As the planets accreted from the impact of smaller bodies, impacts released heat. Smaller bodies had less mass to retain heat and therefore cooled more quickly than larger ones. Gravitational separation of a dense core would also release heat, but separation of a core is less likely in a body with weak gravity and a cool interior. Small bodies have smaller amounts of radioactive material for radiogenic heating and smaller thicknesses of insulating mass to retain the heat that is produced. Larger bodies of the inner solar system show considerably more internal activity than smaller ones.

Spectroscopic evidence indicates that some large asteroids have basaltic surfaces, which means that bodies a few hundred kilometers in diameter have undergone enough internal heating at some time in their history to melt rocks. The Moon is believed to have a very small core, probably rich in iron or iron sulfide. Mercury is unusually dense for a small planet and probably has a core with a radius three-fourths that of the planet itself. Although both Mercury and the Moon are covered with vast volcanic plains, neither has experienced significant volcanic or tectonic activity in the last 3 billion years, perhaps because their interiors are too cool and rigid to permit much convection or melting.

Venus is only slightly smaller than Earth, and its interior is probably similar to Earth's. The surface of Venus shows many manifestations of intense internal activity such as fracturing and folding of the crust and widespread volcanic activity. Nevertheless, Venus lacks the sharp division between two kinds of crust that is so evident on Earth. Also, Venus lacks obvious topographic features typical of plate tectonics on Earth: ridges with crustal spreading, fracture zones, or trenches. One possible reason for its different geology may be that Venus has a thin, flexible crust that results from its very high surface temperature (650 kelvins), rather than thick, rigid crustal plates, as on Earth.

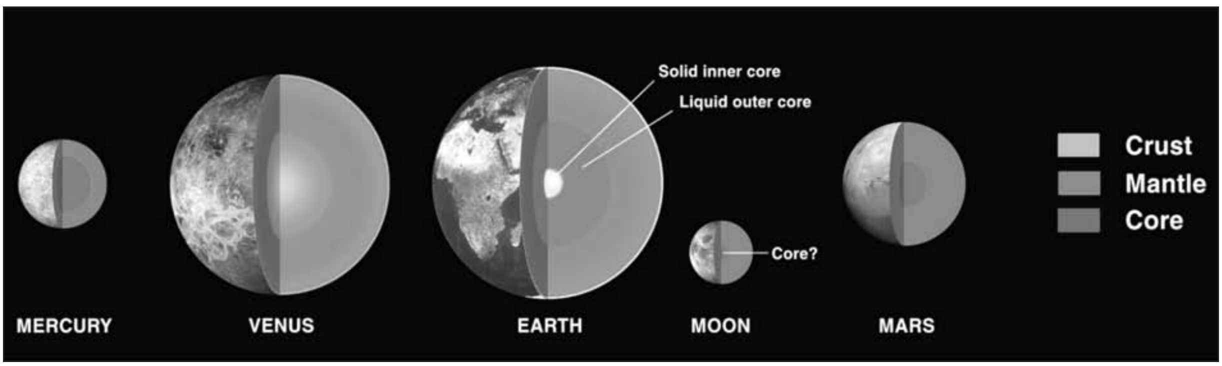
Mars is believed to have a small core of iron or iron sulfide and shows considerably more evidence of recent dynamic internal activity than Mercury or the Moon. A gigantic rift valley, Valles Marineris, extends more than 5,000 kilometers on Mars. The Tharsis Ridge is close by, crowned with several enormous volcanoes. The relative lack of craters on these features of Mars indicates that they formed much more recently than the ancient volcanic plains on Mercury or Venus, probably about one billion years ago. Nevertheless, the opposite side of Mars consists of ancient cratered terrain that has been inactive since its early history. Mars never developed a global system of crustal plates as had Earth, probably because the planet cooled enough for a very thick, rigid lithosphere to form.

In the outer solar system, water ice and other volatile materials (collectively termed "ices") are major constituents of planetary bodies. Water ice evaporates quickly in a vacuum even from the solid state. Only at 175 kelvins or a little sunward of Jupiter was it cold enough for ice to form in a vacuum and survive billions of years. At 150 kelvins, methane hydrate (a solid mixture of methane and water) condensed, followed by ammonia hydrate at 120 kelvins, and argon and pure methane at about 65 kelvins. These temperatures are comparable to the surface temperatures on the satellites of Jupiter (135 kelvins), Saturn (100 kelvins), Uranus (85 kelvins), and Neptune (50 kelvins).

Satellites of all the large outer planets, and perhaps the dwarf planet Pluto and its satellite Charon as well, consist of rocky cores surrounded by mantles of ices. One of the great surprises of planetary exploration has been the degree of internal activity that occurs on objects made largely of ices. A rocky core need not be very large to generate enough radiogenic heat to warm water ice to near or above the melting point, and some larger bodies, especially the large satellites of Jupiter, may have liquid water or icy slush within their mantles. A number of satellites in the outer solar system show evidence that ice has flowed upward from the interior to the surface. This process has been called "ice volcanism," but it is more nearly akin to vertical glacial flow. Water ice undergoes a remarkable series of changes in crystal structure at high pressure, and these high-pressure forms of ice are present undoubtedly in the interiors of large satellites in the outer solar system.

The icy satellites of the outer solar system circle very massive planets and experience a force that is only of minor importance on Earth: tidal forces. Tidal forces arise because the gravitational attraction between two bodies is greater on their near sides than their far sides. Tidal forces distort bodies into ellipsoidal shapes whose long axes point toward the other attracting body. The Moon exerts only a small effect on the solid Earth and a more noticeable effect on the oceans, but Earth's tidal force has locked the Moon's rotation so that it always shows the same face to Earth. Tidal stresses also produce very small "moonquakes" when the Moon is nearest to Earth in its orbit.

Almost all satellites in the solar system are similarly locked to their planets. If a planet has several large satellites, the satellites also exert tidal forces that tend to twist the satellites into



Interiors of rocky terrestrial planets. (Lunar and Planetary Institute) alignment with one another. Under the right conditions, the conflicting tidal forces can generate a significant amount of internal frictional heat. The most spectacular example is Io, Jupiter's innermost satellite. Frictional heating from tidal forces makes Io's interior hot enough to cause active volcanism, despite the fact that Io is smaller than Earth's moon. The ultimate source of energy for this heat is the orbital motion of Jupiter's satellites and the rotation of Jupiter. Heating of Io is but a negligible energy drain on these bodies. Other satellites, notably Europa, another satellite of Jupiter, the Saturnian satellite Enceladus, and several satellites of Uranus have extensive fracture networks also related to tidal stresses.

Jupiter, Saturn, Uranus, and Neptune, the so-called gas giants, probably formed by accumulating rock and ice cores several times as massive as Earth. These large cores were massive enough to attract and hold large amounts of gases. Jupiter and Saturn, which were massive enough to retain essentially all of their gases, are similar to the Sun in composition. Uranus and Neptune did not become massive enough to attract or retain as much hydrogen and are made mostly of denser gases such as ammonia or methane. The deep interiors of all of these planets are very dense gases or fluids. In Jupiter and Saturn, pressures become great enough that hydrogen is believed to have metallic properties, a form of matter predicted in theory and only created in small systems in the laboratory. Fluid movements in the interiors of the gas giants are believed to account for their magnetic fields, in the same manner that fluid motions in Earth's core produce a magnetic field.

METHODS OF STUDY

The deepest mine on Earth is 4 kilometers deep, and the deepest well ever drilled is 12 kilometers deep, about one one-thousandth the diameter of Earth. Direct sampling of the Earth's interior has not penetrated through the crust, much less reached the core. Geologic processes have brought materials from the deep crust and upper mantle to the surface, from depths perhaps greater than 100 kilometers. Knowledge of Earth's deep interior is derived from a variety of indirect techniques, therefore, and knowledge of the interiors of other planets is based in part on knowledge of Earth's interior.

Some information about Earth's interior can be derived from the surface. Most rocks have densities of 2.7 to 3.0 grams per cubic centimeter. The size of Earth and hence its volume are known from mapping. Earth's mass is known from its gravitational attraction. The average density of Earth, 5.5 grams per cubic centimeter, implies that there is dense material deep in the interior. Another observation supports this conclusion. All rotating objects, including Earth, can be described by a physical quantity called moment of inertia. Moment of inertia plays much the same role in equations of rotational motion that mass plays in equations of linear motion. Earth's moment of inertia can be found from observations of gravitational effects of the Moon on Earth. It does not match the moment of inertia of a uniform body with the size and mass of Earth; instead, Earth's moment of inertia indicates that much of Earth's mass is concentrated near its center. Similar studies applied to other planets also provide clues to their internal structure.

Samples of Earth's mantle are found in two geologic settings: as parts of ophiolites (fragments of oceanic crust incorporated into mountain ranges) and as inclusions in kimberlite pipes, volcano-like vents that appear to have brought rocks (and occasionally diamonds) from the mantle to the surface with great speed and violence. The composition of mantle rocks—predominantly iron and magnesium silicates—agrees with the theory that Earth is largely composed of chondritic raw material and their mechanical properties match the properties of the upper mantle as deduced from seismic studies.

One measure of solar-system bodies can provide great insight into their composition, even for bodies not sampled directly: bulk density, which is the mass of a body divided by its volume. The volume of a body can be computed readily once the diameter is known, and diameter can be obtained with high precision with spacecraft imagery. The same spacecraft that obtains imagery can also provide an accurate mass determination, through the gravitational effect of the body on the path of the spacecraft.

Bodies with bulk densities of 3 to 4 grams per cubic centimeter, such as the Moon, Mars, and Io, are probably composed mostly of silicates of the type found in Earth's mantle.

The only solid that is abundant in the solar system and less dense than silicate rocks is water ice, with a density of 0.9 gram per cubic centimeter. Small solid bodies with densities of 1 to 3 grams per cubic centimeter are very likely made of varying proportions of silicate rock and water ice. Most of the satellites in the outer solar system are of this composition. The very large outer planets also have low bulk densities (Saturn has a bulk density of only 0.7 gram per cubic centimeter). These planets are known, from spectroscopic evidence as well as direct imaging by spacecraft, to be composed mostly of dense gases, probably with solid cores.

CONTEXT

Until seismic methods became available, the only way of knowing anything of Earth's interior was by analyzing its exterior properties, such as mass, shape, and gravitational attraction. During the eighteenth and early nineteenth centuries, accurate determination of the shape of Earth was at the foremost frontiers of research in science. Many now-indispensable advanced mathematical methods were stimulated by study of Earth's shape and gravitation.

Seismic studies of Earth's interior began in the late nineteenth century. John Milne, an English engineer working in Japan, refined the crude seismographs then in existence to new heights of sensitivity. By the 1880's, seismographs were capable of detecting earthquakes at distances of thousands of kilometers. In 1896, Richard Dixon Oldham argued from astronomical and geologic evidence that Earth must have a dense iron core. By 1906, he presented evidence for a core by showing that seismic waves passing through the center of Earth traveled more slowly than they did in the mantle. In 1909, the Yugoslavian seismologist Andrija Mohorovičić noted that seismic waves for nearby earthquakes showed an abrupt change in velocity beyond a certain distance, which he interpreted as evidence for a thin outer crust. By 1912, Beno Gutenberg had determined accurately the dimensions of the core. In 1936, the Danish seismologist Inge Lehmann presented evidence for the existence of the inner core. Until the 1920's, most seismographs were mechanical, curious blends of massive weights and delicate lever mechanisms to obtain enough sensitivity and magnification to record faint signals. Since that time, electronic signal amplification has vastly improved the sensitivity of seismographs.

One of the great conceptual advances in the understanding of Earth's interior was the hypothesis of continental drift, proposed by the German meteorologist Alfred Lothar Wegener in 1912. Beginning in the late 1950's, evidence began to accumulate that the active processes for crustal motion on Earth actually were concentrated in the ocean basins. Between 1960 and 1975, a modified concept of continental drift, or plate tectonics, became the generally accepted view among geologists.

Study of the interiors of other planets was not possible until spacecraft could provide accurate dimensions and masses for these objects, as well as views of their surfaces that might provide clues to internal processes. The only body besides Earth to be investigated seismically is the Moon. Apollo astronauts placed seismographs on the Moon between 1969 and 1972. Some spent third stages of the Saturn V rocket and some discarded ascent stage of the Lunar Module were purposely impacted on the Moon to create seismic events of known magnitude for surface experiments to detect. These impacts helped to calibrate the experiments, and also to decipher subsurface structure of the Moon. Seismic waves caused by natural meteor impacts have also helped probe the lunar interior.

One of the most dramatic insights into planetary interiors was provided by American scientists Stanton Peale, Patrick Cassen, and Ray Reynolds. They calculated the amount of tidal heating that might occur on Io and found that Io could get hot enough to sustain volcanic activity. They published this hypothesis in the journal *Science* less than a month before the Voyager spacecraft returned the first photographs of volcanic eruptions on Io, the first active volcanism ever seen outside Earth.

Steven I. Dutch

FURTHER READING

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. A survey of the solar system in the light of Voyager encounters with Uranus and Neptune. Abundantly illustrated and written at the introductory college reading level.

Blondel, Phillippe, and John Mason, eds. *Solar System Update*. New York: Springer, 2006. A series of reviews of selected topics within solar-system science. For the professional astronomer.

Esposito, Larry W., Ellen R. Stofan, and Thomas E. Cravens, eds. *Exploring Venus as a Terrestrial Planet*. New York: American Geophysical Union, 2007. A collection of articles covering all major areas of Venus planetary research. Technical.

Faure, Gunter, and Teresa M. Mensing. *Introduction to Planetary Science: The Geological Perspective*. New York: Springer, 2007. Designed for college students majoring in Earth sciences, this textbook provides an application of general principles and subject material to bodies throughout the solar system. Excellent for comparative planetology.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college text on astronomy, somewhat more advanced than many introductory texts, with a wealth of detail and excellent diagrams. Chapters 6 through 16 describe the solar system. Comes with a CD-ROM.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science. Sepate chapters cover all the planets and their complex systems. Accessible to the general reader and suited for use in a college course at the introductory level. Heavily illustrated.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Includes discussion of the gas giants' interiors. Filled with figures and photographs. Suitable as a textbook for upper-level college courses in planetary science.

Leverington, David. *Babylon to Voyager and Beyond: A History of Planetary Astronomy*. New York: Cambridge University Press, 2003. A historical approach to planetary science, concluding with a summary of spacecraft discoveries. Suitable for general readers and the astronomy community. Heavily illustrated.

Lewis, John S. *Physics and Chemistry of the Solar System*. 2d ed. San Diego, Calif.: Academic Press, 2004. Suitable for an undergraduate course in planetary atmospheres, but accessible to the general reader with a technical background. Includes discussion of the compositions of planetary interiors.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. Planetary atmospheres are treated as important physical features of the various members of the Sun's family. They are discussed individually in the context of what is known about each planet's characteristics and with regard to theories about their evolution and the evolution of the entire solar system. Geared for the undergraduate college student.

Squyres, Steve. *Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet*. New York: Hyperion, 2006. Written by the principal investigator for the Mars Exploration Rovers Spirit and Opportunity, this fascinating book provides a general audience with a behind-the-scenes look at how robotic missions to the planets are planned, funded, developed, and flown. Discusses surface geology, which is influenced by activity in Mars's interior in ages past.

See also: Lunar Interior; Neptune's Interior; Saturn's Interior; Uranus's Interior; Venus's Volcanoes.

Planetary Magnetospheres

Category: Planets and Planetology

Six planets in the solar system possess magnetic fields apparently generated through internal processes. These fields extend some distance into surrounding space and interact with charged particles streaming from the Sun. Electromagnetic interactions occur in the upper atmospheres of these planets, causing phenomena including auroral displays.

OVERVIEW

Planetary magnetospheres are usually created by magnetic force fields generated from the deep interiors of the planets. Six planets of the solar system—Mercury, Earth, Jupiter, Saturn, Uranus, and Neptune—exhibit magnetospheres of differing magnitudes, apparently produced essentially in this way. The mechanics of the generating processes are not well understood. They differ somewhat among the planets, but the effect is as if strong bar magnets lay inside each of the affected planets. Lines of magnetic force radiating from the poles of each of these magnetic fields arc outward into surrounding space. The stronger ones link with corresponding force lines from the opposite pole to create a roughly spherical magnetic field extending several planetary radii into space around the planet. These three-dimensional magnetic structures are complex and dynamic regions. They contain plasmas of varying temperature and density, through which high-energy particles move. Venus and Mars also have magnetospheres, but they are the result of induced magnetism created by the action of the solar wind impacting directly on these planets' atmospheres rather than by internal processes near their cores.

The importance of planetary magnetospheres lies in the ways they control the behavior of electrically charged particles, such as electrons, protons, and atomic and molecular ions. In the inner solar system, the heliosphere contains several charged particles per cubic centimeter on average, and although this is incredibly tenuous by Earth's standards, it amounts to an important factor in the overall environment in which the planets exist.

Most of these charged particles are solar wind particles streaming outward from the Sun at speeds averaging 400 kilometers per second. They derive from the solar corona, a region outside the visible disk of the Sun, where the solar plasma is extremely rarefied and incredibly hot—about 1 million kelvins. Here, protons have such kinetic energy that they escape easily from the Sun's gravity, with the result that they create a steady "plasma wind" that blows out through the solar system. In addition, the complicated and energetic dynamics of the Sun's photosphere and chromosphere create a variety of circumstances in which bursts of charged particles from within are hurled out into the solar system, similar to water from a lawn sprinkler spewing out in a circular pattern.

Were it not for the effects of the solar wind, planetary magnetospheres would be completely spherical in shape. The pressure exerted by the onrushing plasma distorts the magnetic field lines, and therefore the magnetospheres, into teardrop shapes. In the sunward direction, the boundary of a magnetosphere occurs where the internal pressure of the planet's magnetic field balances the pressure of the solar wind. In the direction away from the Sun, the magnetosphere stretches out into a very elongate region called the magnetotail.

The actual size of a magnetosphere varies somewhat from day to day, even from hour to hour, depending on the intensity of the solar wind. Bursts of intense solar activity result in "gusts" in the solar wind which, when they encounter the planetary magnetic fields, compress the magnetospheres. Magnetospheres expand again when the solar wind diminishes. The farther out in the solar system a planet lies, the less is the ambient pressure exerted by the solar wind. This fact explains why the magnetospheres of Jupiter and Saturn are so large that they dwarf the Sun.

Magnetic field lines rotate with the planet's interior, from which they originate. As a result, the entire magnetosphere rotates with what is called "rigid body motion," behaving as if it were a rigid extension of the planet's core rather than a tenuous fluid on the fringes of the planet's sphere of influence. The energy expended to force the magnetosphere to rotate with the interior of the planet is small but significant in the long run and gradually drains a planet of rotational momentum.

At the outer boundary of a planet's magnetosphere, the solar wind breaks like sea waves on a reef. Much of the oncoming material is deflected and flows around the planet in the teardrop-shaped magnetosheath, but not without much turbulence and the formation of a shock wave identical to the wave that forms around aircraft traveling at supersonic speeds. This is known as the bow shock. A small percentage of the particles penetrate the magnetopause (the outer boundary of the magnetosphere) and enter the planet's magnetic field.

In addition to the charged particles acquired from the solar wind, the magnetospheres contain ions and electrons derived from the planet to which they belong and any satellites or rings that may orbit the planet within its magnetosphere. These particles may have drifted from the main body of the planet's atmosphere through natural kinetic activity, or may have been knocked free by collisions with high-energy particles impacting on the upper atmosphere. Once within the magnetosphere, charged particles migrate to particular regions, depending on their kinetic energy, their electrical charge, and other physical factors. Planetary magnetospheres differ with respect to some of these regions. One that appears to be similar in all is the plasma sheet that lies at the center of the magnetotail. The tail consists of two lobes of opposing polarity, between which is a narrow zone where the magnetic field strength is nil. This narrow but elongate region becomes a deep "pocket" in which charged particles can collect. Periodic disturbances such as intense solar activity produce temporary changes in the size and shape of the magnetotail. These can result in sudden injections of large numbers of these stored particles back into the polarized regions of the magnetosphere. Such injections become the source of the particles that create the aurora borealis (northern lights) and aurora australis (southern lights) long known to be associated with solar activity.

METHODS OF STUDY

Information about Mercury's interior structure and magnetosphere depended upon observations and measurements acquired by instruments aboard the American spacecraft Mariner 10, which reached the planet on March 29, 1974. Mariner 10 obtained measurements that enabled researchers to establish with considerable accuracy the size and shape of the magnetosphere and thereby infer the strength of the magnetic field. The field is inclined 12° to Mercury's rotational axis; however, it is much smaller and has only about 1 percent the strength of Earth's field. In two other important respects, Mercury's magnetosphere differs from the magnetospheres of Earth and the Jovian planets. It does not possess any radiation belts, and because Mercury lacks any appreciable atmosphere, none of the interactions between a magnetosphere and an ionosphere is present. More information about Mercury's magnetic field will be obtained when the Mercury Surface Space Environment, Geochemistry, and Ranging (MESSENGER) spacecraft enters orbit about the planet, planned for 2011.

Venus lacks an internal magnetic field on a planetary scale; there may be residual magnetism in its rocks. This is believed to be the result of the fact that the planet spins so slowly on its axis that no dynamo effect occurs at the planet's core. Nevertheless, Venus exhibits a weak ionospheric magnetic field created by an induced magnetism that results from the interaction of the solar wind with the ionosphere. Lacing through the ionosphere are a number of unique magnetic structures called flux ropes, which are long, twisted lines of magnetic force. Within the flux ropes, the plasma temperature and particle densities are high and a wide variety of ions are found.

The Earth's magnetosphere extends about 60,000 kilometers in a sunward direction. The magnetotail stretches out 1,000 radii in the opposite direction. Clear evidence that the magnetic axis is inclined 11° to Earth's rotational axis is seen in the locations of the magnetic poles; the magnetic north pole lies in northernmost Canada and the magnetic south pole is found on the edge of Antarctica. The center of the Earth's magnetic field also lies about 500 kilometers away from the geometric center of Earth.

Many of the charged particles trapped within the Earth's magnetosphere are concentrated in regions called the plasmasphere and the Van Allen radiation belts. The plasmasphere begins just beyond the ionosphere and grades outward into the Van Allen radiation belts. There are two belts: The inner belt extends radially from about 1,000 to 10,000 kilometers, and the outer belt from 25,000 to 60,000 kilometers. The Van Allen radiation belts surround the planet like two concentric doughnuts, encasing the equatorial and temperate latitudes but leaving the poles relatively exposed. Electrons, the least massive of the charged particles and most easily deflected by magnetic fields, spiral around the lines of force in these radiation belts as they travel back and forth between the magnetic poles at speeds of thousands of meters per second.

Poleward of the Van Allen radiation belts is a region of the magnetosphere that does not trap charged particles but funnels them down toward the atmosphere. The impact of the charged particles upon atoms and molecules in the polar ionosphere raises the temperature of this region to more than 1,650 kelvins, causing atmospheric gases to fluoresce and create what is known as the northern lights and southern lights. Seen from a perspective high above the planet, auroral activity appears as rings of glowing light girdling the polar ionosphere, like halos over the Arctic and Antarctic regions.

Spacecraft have detected only a faint trace of a magnetic field around Mars. Its strength is less than one-tenth of 1 percent that of the Earth. This is not considered surprising, since the planet lacks a liquid core and cannot generate a field internally. Evidently, what has been detected is similar to the induced magnetism in the ionosphere of Venus, but the Martian atmosphere is much thinner, making even this effect weak. In situ examination of rocks on the surface by the Mars Exploration Rovers Spirit and Opportunity found residual magnetism in certain rocks.

Detection of sharp bursts of radio energy from Jupiter in 1955 was the first indication that this planet might possess a magnetosphere. Further study—particularly of data provided by Pioneers 10 and 11, Voyagers 1 and 2, and Galileo—confirmed the existence of a huge magnetosphere and provided many insights into its features. Jupiter's magnetic axis is inclined about 10° to the axis of rotation, and its field is nineteen thousand times stronger than Earth's. The magnetosphere extends in the sunward direction an average distance of 3 million kilometers, while the magnetotail stretches beyond the orbit of Saturn. The polarity of the field is inverted, as compared with Earth's (that is, the magnetic north pole lies near the planet's geographic south pole). In the huge sea of plasma surrounding Jupiter are protons, electrons, ions, and neutral atoms from three distinct sources: the solar wind, the planet's atmosphere, and the surfaces of some of its satellites. The third source contributes the heavy ions that are present through the process of sputtering. There is no correspondence for this component in Earth's Van Allen radiation belts.

Io, the innermost of the Galilean satellites, is the site of several active volcanoes. Each of these sites discharges as much as 100,000 tons of sulfur or sulfur dioxide per second. Most of this material rains back down on the surface, but a very small fraction of it (0.01 percent) escapes from Io's gravity and enters the Jovian magnetosphere. There it is quickly broken down by sunlight and fast-moving solar particles whose impacts cause the formation of heavy ions of oxygen, sodium, and sulfur. A doughnut-shaped volume within the Jovian magnetosphere, known as the Io Torus, contains a concentration of these heavy ions. Io's movement through this charged environment generates an intense electrical current that follows a huge arcing path between Io and Jupiter, known as the Io flux tube. As Io orbits above the portion of Jupiter facing toward Earth, the magnetic "footprint" of this flux tube faces Earth and unusual bursts of radio energy are detected as this concentration of particles interacts with the Jovian atmosphere.

The existence of Saturn's magnetic field was not detected until Pioneer 11 approached the planet in 1979. The field is inclined exactly with the planet's axis of rotation and is one thousand times stronger than Earth's but nineteen times weaker than Jupiter's. Like Jupiter, the magnetic polarity is inverted as compared with Earth's. On the sunward side, Saturn's magnetosphere extends an average of 1 million kilometers from the planet. Measurement of the field strength at the cloud tops shows the magnetic north pole to be stronger than the south pole, suggesting that the magnetic center of Saturn lies about 2,400 kilometers north of its geometric center. Titan orbits Saturn at a mean distance of 1,221,400 kilometers, passing out of the magnetosphere as it comes around on the sunward side, only to plunge back into the magnetotail as it continues its journey. Titan's significant atmosphere (1.6 times as dense as Earth's) is the source of heavy ions of nitrogen and other molecules, along with a significant amount of neutral hydrogen. The inumerable particles in Saturn's rings, however, have been shown to be excellent absorbers of the charged particles in the magnetosphere. Absorption of electrons by the rings is so complete that a region beginning at the outermost of the three main rings (the A ring) and extending all the way to the planet's surface is the most radiation-free zone in the solar system.

Uranus's magnetosphere is generally similar in size and strength to Saturn's, but not in composition. It contains little other than protons and electrons derived from hydrogen escaping from the planet's atmosphere. The inner magnetosphere includes an extensive corona of protons similar to Earth's Van Allen radiation belts. The absence of significant numbers of heavy ions in the magnetosphere may be attributable to the fact that the energies of the protons are lower than is the case with Jupiter and Saturn, consequently limiting their ability to sputter heavy ions off the surfaces of the satellites and rings. The magnetic axis is inclined 60° to the planet's rotational axis, but since Uranus's rotational axis is out of phase radically with the rotational axis of the other planets, this places the magnetosphere in a fairly normal orientation to the plane of the solar system. Its polarity matches that of Jupiter and Saturn. Not only is the magnetic axis steeply inclined to the rotational axis, but also it is offset a distance equal to about one-third of the planet's radius from the geometric center of Uranus. The radical inclination of the two axes gives rise to some unusual dynamics in the magnetosphere. As the magnetosphere rotates with the planet's interior, it creates a corkscrewing effect in the plasma of the magnetotail.

Neptune's magnetosphere was first revealed to the Voyager 2 spacecraft and made its presence known through weak nonthermal radiation with a sixteen-hour repetitive pattern that scientists believe is linked to the planet's period of rotation. The field shares the same polarity as the fields of Earth and Mercury. Two unusual features reminiscent of Uranus are the magnetosphere's rakish tilt of 47° to the rotational axis and the location of the center of the field at a point 14,000 kilometers away from the geometric center of the planet. Particle density within the magnetosphere is the lowest of any of the planets, with no more than 1.4 protons and heavy ions per cubic centimeter present. Particle densities at Uranus and Jupiter are three and three thousand times greater, respectively.

CONTEXT

At the midpoint of the twentieth century, scientists believed that the Earth's upper atmosphere trailed away to an utter and complete vacuum only a few hundred kilometers above the surface. Today, it is known that the atmosphere grades gradually into the huge magnetosphere, within which are pockets of surprisingly concentrated plasma as well as vast regions where little exists. The Earth's magnetosphere and those of other planets are, in turn, embedded in the much larger solar magnetosphere, or heliosphere. This complex mesh of electrical and magnetic fields, electric currents, and particle flows becomes a conduit by which electromagnetic disturbances occurring on the Sun are conveyed directly to the planet's ionosphere, and even to its surface.

Some consequences of this linkage are relatively well known, such as the fact that high levels of solar activity interfere temporarily with shortwave radio broadcasts by altering the upper atmosphere's reflectivity to radio waves. Solar disturbances also interfere with transmission of electric power and communications over long-distance cable systems. It is necessary to engineer such systems with features that prevent the disruption of service when the Earth's magnetosphere is in turmoil because of solar activity.

There is growing physical evidence suggesting a connection between solar activity and terrestrial weather. Historical evidence links a prolonged period of solar inactivity from 1645 to 1715 C.E. (known as the Maunder minimum) to severe climate changes. The mechanism for these linkages is not yet clear, but there is no doubt that the ionosphere responds profoundly to changes in the magnetosphere. Some scientists believe that changes in the ionosphere are conveyed to the lower atmosphere, where weather occurs, through an electrical current that flows between the upper atmosphere and the ground. This current, called the global circuit, is involved intimately in the thousands of thunderstorms occurring over Earth every day. It has also been suggested that there may be a connection between solar activity and geophysical events such as earthquakes.

Records in the Earth's rocks, particularly the iron-rich rocks of the seafloor, prove conclusively that the Earth's magnetic field has reversed polarity on many occasions. There is no cyclical pattern evident in the intervals between reversals, and their cause is not yet known. It has been speculated that major meteor impacts may cause the poles to "flip-flop." However, the phenomenon may also be a response to changes in the Sun's polarity, transmitted through the charged heliosphere. Whether there are any environmental consequences to these polarity changes and whether they can be predicted still need to be determined.

On a cosmological scale, radio astronomy and space-age probes have demonstrated that the universe contains many vast structures of plasma organized around magnetic force fields. They are detectable by the radio energy they give off and are found to surround other stars and even entire galaxies. In fact, such structures may occur as frequently as the much more familiar gravitationally organized ones. The Earth's magnetosphere is, therefore, an important laboratory in which the behavior of ionized matter can be studied. It is reasonable to believe that as scientists understand plasmas and magnetospheres better, new insights will be gained, many of which may have important scientific and practical implications.

Richard S. Knapp

FURTHER READING

Burgess, Eric. *Venus: An Errant Twin*. New York: Columbia University Press, 1985. A comprehensive physical description of Venus for nontechnical readers. The subject of Venus's magnetosphere is presented first in terms of the information gleaned by each of the many space probes to reach the planet, and then in the context of a definitive synthesis of what is presently known about the physical conditions in, on, and surrounding the planet. Suitable for high school and college readers with a good general science background.

Cross, Charles A., and Patrick Moore. *The Atlas of Mercury*. New York: Crown, 1977. Discusses the surface features of Mercury. The final chapter is devoted to the atmosphere, the interior, and the magnetosphere. Excellent diagrams and graphs will benefit readers not familiar with concepts of magnetospheres and the techniques for measuring them from spacecraft. The brief text is clear, well researched, and stimulating.

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001. A challenging and thorough text for students of planetary geology. An excellent reference for the readers with a strong science background.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college text on astronomy, somewhat more advanced than many introductory texts, with a wealth of detail and excellent diagrams. Chapters 6 through 16 describe the solar system. Comes with a CD-ROM.

Halliday, David, Robert Resnick, and Jearl Walker. *Fundamentals of Physics, Extended*. 9th ed. New York: Wiley, 2007. This textbook has taught millions of college students the fundamentals of physics. Its sections on magnetism are particularly strong. Even for those not familiar with basic calculus, there is much to be found to explain the nature of charged-particle behavior in electromagnetic fields.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune, including their magnetic fields and magnetospheres. Filled with figures and photographs and accessible to a general audience.

Lewis, John S. *Physics and Chemistry of the Solar System*. 2d ed. San Diego, Calif.: Academic Press, 2004. Suitable for an undergraduate course in planetary atmospheres, but accessible to general readers with a technical background, this text includes coverage of planetary magnetic fields.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar system astronomy suitable for an introductory college course as well as nonspecialists. Filled with supplemental learning aids and solved student exercises. A Web site is available for educator support.

Moldwin, Mark. *An Introduction to Space Weather*. Cambridge, England: Cambridge University Press, 2008. This text introduces space weather—the influence the Sun has on Earth's space environment—to the nonscientist. Discusses both the scientific aspects of space weather and issues of technological and societal import.

Serway, Raymond A., et al. *College Physics*. 7th ed. New York: Brooks/Cole, 2005. Designed for introductory college physics courses, this book covers the basics of magnetic fields. Filled with sample problems. Comes with PhysicsNow, an online teacher/student resource.

See also: Auroras; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth-Sun Relations; Jupiter's Magnetic Field and Radiation Belts; Saturn's Magnetic Field; Solar Wind; Uranus's Magnetic Field; Van Allen Radiation Belts.

Planetary Orbits

Category: Planets and Planetology

Planets in the solar system revolve around the Sun in elliptical orbits at speeds that vary with distance from the Sun. Laws that govern these motions were first deduced by Johannes Kepler and later quantified by Sir Isaac Newton.

OVERVIEW

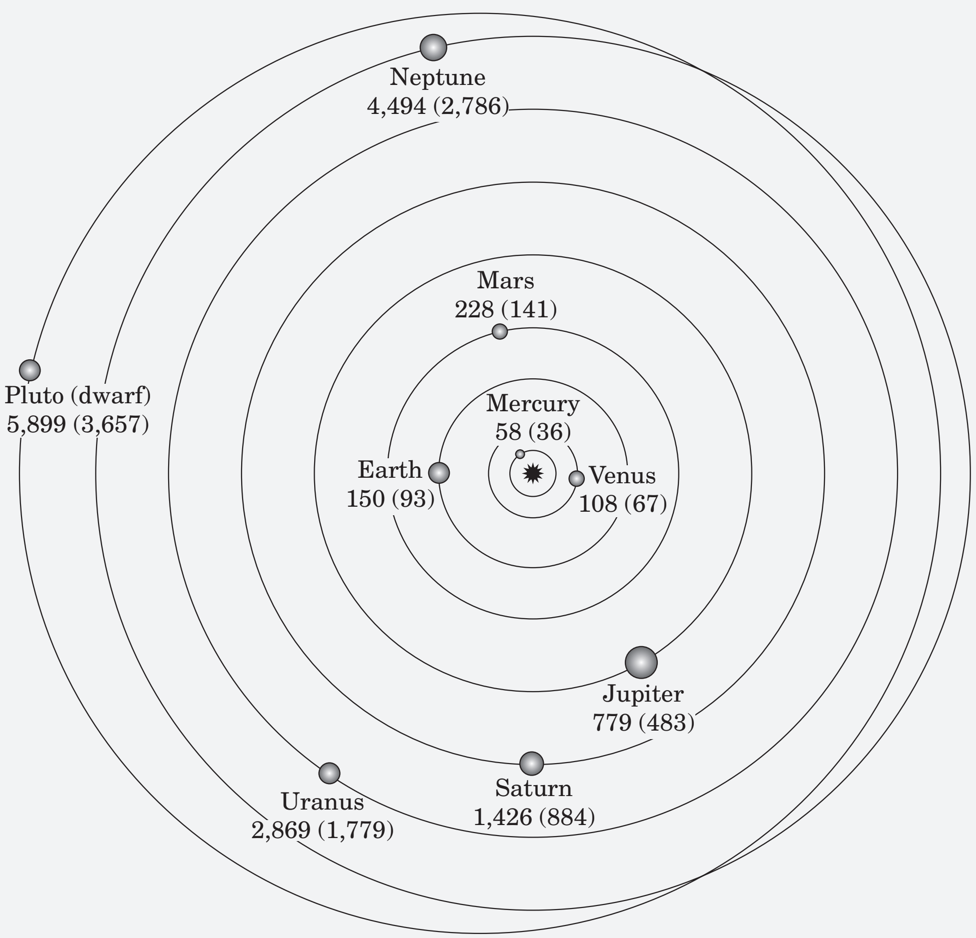
Planets in the solar system move around the Sun in elliptical orbits. Those whose orbits are closest to the Sun move more rapidly than those that are farther away. These simple, universally accepted observations form the basis of the knowledge of planetary motions. Gravity—the force that causes apples to fall from trees and keeps humans firmly planted on Earth's surface—plays the central role in the mechanics of planetary motions.

A simple experiment illustrates the energy relationships inherent in orbiting bodies. If a person attaches a string to a small rubber ball and the ball is swung around the person's head in a horizontal circle, the tension along the string that holds the ball in its "orbit" is analogous to the Sun's gravity pulling on a bound planet. The English astronomer and mathematician Sir Isaac Newton (1642-1727) explained how the force of gravity affects planetary motion. Newton proved in his laws of motion that once an object is in a straight-line motion, it will continue on that course with no further input of energy (law of inertia) unless its motion is perturbed by an unbalanced force. In the case of planets, this force is provided by the gravitational attraction of the Sun (or more massive planet, in the case of a satellite). Depending upon the magnitude of the orbiting body's "kinetic energy" (energy of motion), the body will move in either a circular orbit or, with greater kinetic energy, an elliptical orbit. Kinetic energy counters the attractive force of gravity, thus preventing a planet from falling into the Sun, or the orbiting ball, as stated in the example, from striking the experimenter.

The scientist who first showed that the orbits of the planets are actually ellipses rather than circles was Johannes Kepler (1571-1630). A German mathematician, astronomer, and astrologer, Kepler worked previously as an assistant to the Danish observational astronomer Tycho Brahe. After Brahe's death, Kepler used detailed position measurements of the planet Mars to plot an orbit that was not circular. Up to this time, planetary orbits—including that of the Moon—were believed to be circular in accordance with precepts developed by the Greek philosopher Aristotle.

A circle is the locus of points all the same distance from a given center. An ellipse differs from a circle in being oval-shaped. An ellipse contains two internal, evenly spaced points called foci. It is important to understand how the foci of an elliptical orbit relate to the positions of an orbiting planet and the Sun. This relationship is expressed by Kepler's first law: Each planet moves around the Sun in an orbit whose shape is that of an ellipse, with the Sun at one focus. The other focus is empty. Thus, the Sun is not precisely in the middle of the ellipse but displaced somewhat to the side. The degree of displacement determines the orbit's eccentricity. As a result, planets move between a minimum distance from the Sun in their orbit, called perihelion, and a maximum distance

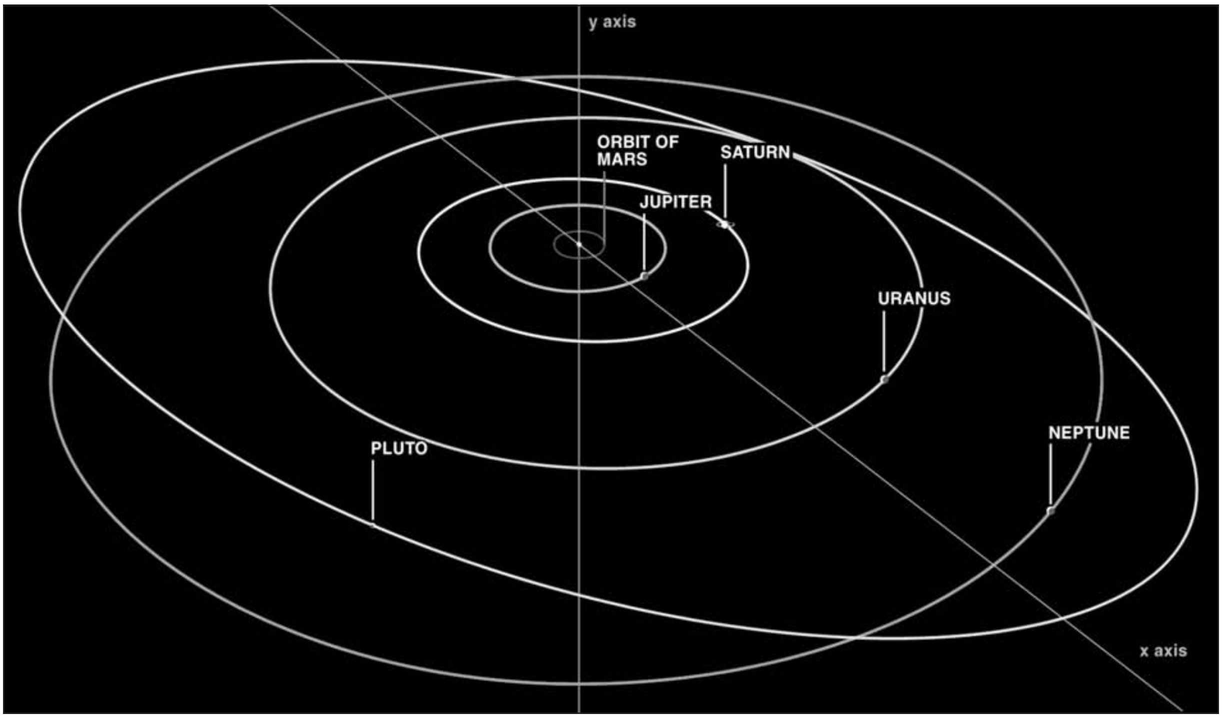
Distances of the Planets from the Sun
in Millions of Kilometers (Miles)



from the Sun, called aphelion. Planetary orbits have this repetitive pattern as a result of the central character of gravity; that is, gravity acts along the line between the gravitationally interacting bodies. The magnitude of the gravitational force follows an inverse square law with regard to its dependence upon distance between the interacting masses. If one doubles the distance between the two objects, their gravitational attraction diminishes not by a factor of two, but by four.

Kepler's second law was actually discovered before his first law. It describes the fact that planets move more slowly when farther away from the Sun (their slowest speed is at aphelion), and they move more rapidly when closer to the Sun (their maximum speed is at perihelion). This observation logically would lend support to the idea that the planet's orbit is anything but circular. The second law states that a straight line joining the planet and the Sun sweeps out equal areas in space in equal intervals of time. Imagine a string attached to a planet at one end and the Sun at the other. When the planet is near aphelion (farthest from the Sun), it moves slowly, so that the triangular sector swept out by the string during a given time will resemble a long, slender piece of pie. In contrast, near perihelion over the same time period, the planet will move farther (because it is going faster), so that the sector swept out by the string resembles a fatter slice of pie. Kepler's second law states that these two pie slices, or triangular sectors—although quite different in radius and opening angle—should have equal areas. This exercise is a mathematical way of stating that planets move more slowly as the Sun-planet distance increases. Planetary orbits obey Kepler's second law of motion as a consequence of conservation of angular momentum.

Kepler's third law, formulated in 1619, was an attempt to quantify the fact that a planet moves more slowly the farther its orbit is from the Sun. His task was to determine a precise mathematical relationship between a planet's average distance from the Sun and its period. Being oval, ellipses have a major axis and a mi-



This "oblique" view of the outer planets' orbits around the Sun makes clear how much closer Mercury, Mars (labeled), Venus, and Earth are to the Sun than are the "gas giants." (Lunar and Planetary Institute) nor axis of different lengths. A line passing through the two foci of the ellipse and ending at both ends of the figure defines the long axis of the ellipse and is known as the major axis. A length equal to one-half the major axis is called the semimajor axis. A line perpendicular to the major axis passing halfway between the two foci of the ellipse is the minor axis. A length equal to one-half the minor axis is called the semiminor axis. A planet's mean distance is half the sum of the perihelion and aphelion distances. This is equal to the average distance of a planet from the Sun and also is the value of the semimajor axis. Kepler found that the cube of the mean distance for any planet is equal to the square of that planet's period. This equation is expressed mathematically as $p^2 = r^3$, where p is the planet's period in Earth years and r is the planet's mean distance from the Sun, expressed in terms relative to the Earth's mean distance, 150 million kilometers, or one astronomical unit (AU). If the Earth's mean distance equals 1.0 AU, then Mars's mean distance is 1.5 AU, Venus's mean distance is 0.72 AU, and so on. Planetary orbits obey this third Keplerian law of motion as a result of the central character of gravity as well as its inverse-square-law nature of gravity.

Kepler's Laws of Planetary Motion

Johannes Kepler's three laws of motion, articulated in the first years of the seventeenth century, laid the foundation for Sir Isaac Newton's law of universal gravitation.

FIRST LAW: A planet orbits the Sun in an ellipse, with the Sun at one of the two foci.

SECOND LAW: The line joining the planet to the Sun sweeps out equal areas in equal times as the planet travels around the ellipse.

THIRD LAW: The ratio of the squares of the revolutionary periods for two planets is equal to the ratio of the cubes of their semimajor axes. That is, the time it takes a planet to complete its orbit is proportional to the cube of its average distance from the Sun. The farther from the Sun an object is, the more slowly it moves.

Johannes Kepler. (Library of Congress)

Newton later reformulated Kepler's three laws using more sophisticated mathematics than was available to Kepler. Newton's modification of the first law states that each planet has an elliptical orbit with the center of mass between it and the Sun at one focus. The "center of mass" is a point between the two bodies (the Sun and the orbiting planet) where their masses are essentially balanced. Mathematically, it is the point at which the product of mass times length is equal for the two bodies: $M_1 L_1 = M_2 L_2$, where M = mass, L = length from center of mass, and subscripts 1 and 2 referring to bodies 1 and 2. The Sun is such an extremely massive body that its center of mass with any planet lies near the Sun's own center. Therefore, the Sun does lie essentially at a focus of the planetary ellipse, as Kepler stated. Its movement around the center of mass (deep within its interior) is detectable only as a slight wobble. For bodies that are more comparable to one another in terms of mass, such as pairs of stars, these objects actually revolve around a common point that lies between them. Pluto and its similarly sized satellite Charon provide a good example of that effect. Because the masses of these bodies are similar, they revolve around a common point known as a the barycenter.

Newton revised Kepler's second law as follows: Angular momentum in a two-body system is constant when no net external torque is present. This law originally described the fact that planets move more rapidly when they are closer to the Sun compared to when they are farther away. Newton found that all bodies that rotate or move around some center have "angular momentum." This quantity is expressed as a body's mass times its speed times its distance from the center of mass (mvr, where m = mass, v = linear speed, and r = distance from the center of mass). Because angular momentum is constant for any two-body system in the absence of a net external torque, if r becomes greater, v must become smaller to compensate (mass always remains constant). On the other hand, near the center of mass (the Sun, for planets), the distance r is diminished and speed v must increase to compensate. Conservation of angular momentum comes into play when a spinning skater pulls her outstretched arms close to her body, initiating a more rapid spin rate. Physicists and astronomers usually talk about planetary speeds of revolution or more properly angular velocity, which is the linear speed per unit distance from the focus. In such a discussion, angular momentum then involves the product of moment of inertia times angular velocity. There is no net torque acting on the planet revolving about the Sun, so this angular momentum expression is conserved or remains constant. That means that the distance from the Sun squared times angular velocity is an invariant throughout the planet's orbital motion.

Newton's revision of Kepler's third law is especially important. Newton discovered that the sum of the masses of the two bodies times the square of the period is proportional to the cube of the mean distance, which is expressed mathematically as $(M_1 + M_2)P^2 = a^3$. The masses must be expressed as a fraction of the Sun's mass for the calculation to be valid. The immediate consequence of this equation is that astronomers could now use this equation to calculate the masses of distant bodies given information on the mean distance and period of the orbiting bodies. In most instances, the mass of the smaller body (planet or satellite of a planet) may be neglected because that mass is so insignificant compared to the Sun's mass (1.99×10^{30} kilograms, or 332,943 times Earth's mass). Rearranging the equation gives $M_1 = a^3/p^2$. This equation can now be used to calculate the mass M_1 of any central body that has a satellite of mass M_2.

APPLICATIONS

Consequences of Kepler's and Newton's laws of planetary motion and gravity had an impact on the scientific world not only during their own time but to this very day. The results of their work continue to be used by astronomers to solve problems. For example, the flight path of the Apollo astronauts to the Moon and back was calculated using all three of Kepler's laws. The energy required to propel the Saturn 5 rocket on its way and later to orbit the Moon was calculated using Newton's laws of gravity. The same can be said for all interplanetary spacecraft, such as Voyagers 1 and 2, which visited and photographed the outer planets, Jupiter, Saturn, Uranus, and Neptune. The two Voyager probes were assisted in their journeys by using the gravitational attraction of these massive planets to accelerate them toward their next target. Calculating gravity assists involves kinetic energy and gravitational relationships developed by Newton.

One of the most useful of Kepler's laws for planetary astronomers is the third law as modified by Newton. This law allows calculation of the mass of a massive body using data about mean distance and period of one or more of its satellites. It has been used to calculate the masses of all planets that have satellites (which excludes Mercury and Venus). One of the most difficult mass determinations was that for the dwarf planet Pluto and its satellite Charon. These bodies are so far away from Earth that Charon was discovered only in 1977. Its orbital characteristics were determined, with great difficulty, some time later. The similar masses of Charon and Pluto cause them to orbit a center of mass (a barycenter) that lies nearly halfway between them, but the location of that barycenter is somewhat closer to Pluto than it is to Charon. The third law was used to calculate both the mass of Pluto, using data from Charon's orbit, and the mass of Charon, using data for Pluto. These calculations show that both bodies have very low masses and are most likely composed of methane ice.

Another important consequence of the laws of planetary motion involves the survival of life on Earth. One theory suggests that periodic mass-extinction events—such as the demise of the dinosaurs—may have been caused by gigantic impacts of asteroids (rocky planetoids with diameters of less then 1,000 kilometers) or comets (asteroid-sized ice balls) with Earth. In the solar system, most asteroids are concentrated in a belt between Mars and Jupiter, whereas most comets originate in the outer regions of the solar system and beyond. Occasionally, collisions or gravitational perturbations from the massive gas giant planets, such as Jupiter or Saturn, cause asteroids and comets to assume orbits that carry them near Earth. All these bodies have sufficient kinetic energy to resist Earth's gravitational attraction, so that objects that graze Earth's orbit continue by without going into orbit around Earth. This fact explains why Earth and other relatively low-mass planets have few or no satellites (while the gas giants—Jupiter, Saturn, Uranus, and Neptune—have many). Therefore, the bodies that do strike Earth, causing extinctions and making huge craters if they strike land areas, must make a direct hit of a moving Earth. Chances of that occurring on a frequent basis fortunately are rather low, but not zero. Given the billions of years of the history of Earth, it is probable that an occasional body will crash into Earth with catastrophic consequences. The high kinetic energy of these bodies is converted into heat and shock waves upon impact, causing considerable destruction. Newton's laws of gravity and motion play a pivotal role, mostly in determining the trajectories of these dangerous visitors to the inner solar system. By the same token, Newton's laws reveal ways that gravity could be ingeniously used to push possible impacting bodies away from a trajectory that otherwise would have them intersect with Earth, thereby averting a possibly cataclysmic collision.

CONTEXT

The history of science closely parallels the development of astronomy in that the study of heavenly bodies and their relationship to Earth dominated philosophical and religious thinking for millennia. One of the first scientists to study religious thinking and astronomical phenomena seriously was the Greek philosopher Aristotle (384-322 B.C.E.). Unlike most of his contemporaries, Aristotle used some observations to prove his speculations. His major contribution to planetary motion studies was his belief that the natural state of matter is to seek the center of the Earth, which is why objects always fall when released above the Earth. Although erroneous, this and related ideas laid the groundwork for later studies by Galileo and Newton on the effects of gravity. Aristotle also believed, as did many others, that Earth was at the center of the universe. That the Sun and planets revolved around Earth in perfectly circular orbits was advocated first by his great mentor, Plato. Later, Aristarchus (c. 270 B.C.E.), a Greek astronomer, adopted the idea that the Sun is at the center of the known universe. That idea was forgotten until revived nearly two thousand years later by Nicolaus Copernicus, whose "heliocentric" model, published shortly after his death in a volume titled *De revolutionibus orbium coelestium* (1543; *On the Revolutions of the Heavenly Spheres*, 1952; better known as *De revolutionibus*), describes a system in which the planets orbit the Sun in perfect circles. Although not completely accurate, the heliocentric model eventually supplanted the Earth-centered model of Aristotle and other philosophers.

In the middle of the second century C.E., Ptolemy (c. 100-178 C.E.), wrote a text called *Mathēmatikē syntaxis* (c. 150 c.e.; *Almagest*, 1948) in which he summarized all that was known about astronomy up to that time. This book influenced astronomical thinking for the next millennium. It included a model of the solar system that was quite accurate in predicting planetary positions. Using an idea first developed by Apollonius of Perga (c. 240-170 B.C.E.), Ptolemy declared that planets move in perfect circles around Earth. Nevertheless, planets moved simultaneously in smaller circles called "epicycles." These were necessary to explain why the outer planets occasionally seemed to reverse their normally eastern motion relative to the stars and move west. One of the great triumphs of Kepler's laws is that they provide an explanation for retrograde motion (Earth moves faster and overtakes the outer planets). Kepler, and before him Copernicus, laid the foundations for a scientific understanding of planetary motions that broke the hold on thinking imposed by the *Almagest*.

Sir Isaac Newton used ideas developed by Galileo and Kepler to quantify the knowledge of planetary motion and to explain these motions in terms of gravitational forces and kinetic energies. He published his ideas in *Philosophiae Naturalis Principia Mathematica* (1687; *The Mathematical Principles of Natural Philosophy*, 1729; best known as the *Principia*). Although it is now known that Newton's laws do not work well on atomic and subatomic scales (treated in the discipline of quantum mechanics) or in cases where bodies are moving relative to one another at very high speeds close to that of light (which later were addressed by Albert Einstein's relativity theory), Newton's laws work perfectly under everyday conditions on Earth and in the solar system.

Newton's and Kepler's time-tested laws continue to be used by astronomers and other space scientists to make predictions about planetary motions and interactions. Relativity is needed, however, to explain the advance of Mercury's perihelion as it orbits so close to the Sun. Spacecraft deep in the gravitational well of the Sun likewise require calculations involving relativity to maintain them on their proper courses.

John L. Berkley

FURTHER READING

Arny, Thomas T. *Explorations: An Introduction to Astronomy*. 3d ed. New York: McGraw-Hill, 2003. A general astronomy text for the nonscientist. Includes an interactive CD-ROM and is updated with a Web site.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. Filled with color diagrams and photographs, a popular work on solar-system astronomy and planetary exploration through the Mars Pathfinder and Galileo missions. Accessible to the astronomy enthusiast. Provokes excitement in the general reader, who gains an explanation of the need for greater understanding of the universe around us.

Consolmagno, Guy. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994. A text accessible to college-level science majors and general readers alike. Presents explanations using low-level mathematics and also involves integral calculus where required. Demonstrates how the area of planetary science progresses by questioning previous understanding in the light of new observations.

Halliday, David, Robert Resnick, and Jearl Walker. *Fundamentals of Physics, Extended*. 9th ed. New York: Wiley, 2007. This textbook has taught millions of college students the fundamentals of physics. Its sections on Newton's laws of motion are particularly strong, as is the chapter on gravitation, which includes derivations of Kepler's laws of planetary motion. Even for those not familiar with basic calculus, there is much to be gained by studying from this all-encompassing work.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college-level text that is clearly written; most nonspecialists should find this book a rich source of information. Chapter 3, "Celestial Mechanics," discusses the historical development and application of the laws of gravity and motion. Contains detailed black-and-white diagrams and photographs. Tables in the appendix offer comprehensive data on planetary orbital characteristics and other useful information.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find the standard treatise for typical ASTRO 101 classes too low level. Suitable for an audience with varied science and mathematical backgrounds. Covers all topics from solar-system objects to cosmology.

Leverington, David. *Babylon to Voyager and Beyond: A History of Planetary Astronomy*. New York: Cambridge University Press, 2003. Takes a historical approach to planetary science. Heavily illustrated, concluding with a summary of spacecraft discoveries. Suitable for general readers and the astronomy community alike.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar-system astronomy suitable for an introductory college course as well as nonscientists. Filled with supplemental learning aids and solved student exercises. A Web site is available for educator support.

Serway, Raymond A., Jerry S. Faughn, Chris Vullie, and Charles A. Bennet. *College Physics*. 7th ed. New York: Brooks/Cole, 2005. A textbook used at the introductory level in college physics courses, filled with sample problems, including those on laws of motion. Comes with an online teacher/student resource.

Snow, Theodore P. *The Dynamic Universe*. Rev. ed. St. Paul, Minn.: West, 1991. A general introductory text on astronomy. Covers the kinematics and dynamics of planetary motion. Features special inserts, guest editorials, and a list of additional readings at the end of each chapter. College level.

Stephenson, Bruce. *Kepler's Physical Astronomy*. Princeton, N.J.: Springer, 1994. A complete historical account of the search to understand the orbital behavior of planets in the solar system. For both the technical and general reader.

See also: Earth's Rotation; Planetary Orbits: Couplings and Resonances; Planetary Rotation.

Planetary Orbits: Couplings and Resonances

Category: Planets and Planetology

Orbital or rotational (spin) motions of objects are said to be "coupled" or "in resonance" when the relationships between the periods of such motions can be expressed as ratios of small integers such as 1:1, 1:2, 1:3, 2:3, or 3:4. This usually occurs as the result of the gravitational interaction of the objects. Many examples of couplings and resonances occur in the orbital and rotational motions of solar system objects.

OVERVIEW

The orbital or rotational (spin) periods of many solar-system objects have been found to be related by ratios of small integers, such as 1:1, 1:2, 1:3, 2:3, or 3:4. Such relationships in the motions are called couplings or resonances. Usually they have developed over time as the result of gravitational interactions between the objects. Because of their ubiquity, couplings and resonances are thought to have played a major role in shaping the structure of the solar system.

The many couplings and resonances in our solar system can be categorized into two main types. One type, called a spin-orbit resonance, is manifested by a simple ratio between an object's period of spinning (rotating) on its axis and its period of orbiting (revolving) around a more massive body. For example, the time it takes the Moon to rotate on its axis exactly equals the time it takes to revolve around the Earth, a ratio of 1:1. As a result, the Moon always keeps the same side facing the Earth. (Because the Moon's orbit around the Earth is not precisely circular, its orbital speed varies slightly. Consequently it appears to us on Earth as if the Moon rocks a bit from side to side—a motion called libration—so we end up seeing slightly more than half the Moon's surface during one of its orbits.) Another example is that Mercury's period of rotating on its axis is exactly two-thirds of its period of revolving around the Sun, a ratio of 2:3. Thus, Mercury spins three times on its axis during two orbits around the Sun.

The other type of resonance, called an orbital resonance, involves small-integer ratios between the orbital periods of two or more small-mass bodies orbiting around a much more massive object. Such relationships reinforce the gravitational interactions between the resonant objects. Suppose one of the orbiting small-mass bodies is much less massive than the other small-mass body. (Examples include an asteroid and Jupiter as both orbit the Sun, and a ring particle and a satellite as both orbit Saturn.) The repeated gravitational tugs of the more massive orbiting body (Jupiter or the satellite) on the less massive orbiting body (the asteroid or the ring particle) will tend to pull the less massive body away from its resonant orbit, while the more massive orbiting body is little affected. This clearing of resonant orbits produces gaps in belt and ring systems. Such gaps, or divisions, are observed in the asteroid belt (the Kirkwood gaps), located mainly at the resonances with Jupiter, and in Saturn's ring system as a result of resonances with some of Saturn's satellites.

When two or more objects have exactly the same orbital period around a more massive object, they are called coorbital, a special case of orbital resonance. In 1772, Joseph-Louis Lagrange mathematically discovered five points at which coorbital bodies could exist in equilibrium. These points are called the Lagrangian points and are labeled L1 through L5. Three of the points (L1 through L3) are unstable, in that an object displaced slightly from the point will drift farther away. However, the L4 and L5 points are stable, in that an object displaced slightly from the point will remain nearby and oscillate around the equilibrium position. These two stable points are located 60 degrees ahead (L4) and 60 degrees behind (L5) the second-largest body along its orbit around the largest body. The Trojan asteroids (so called because they have been named after the heroes, both Trojan and Greek, of the Trojan War) oscillate around the L4 and L5 points 60 degrees ahead of and behind Jupiter along its orbit around the Sun.

APPLICATIONS

Several types of resonance manifest themselves in the solar system. In all cases, gravity provides the coupling force, although the way gravity is applied to cause the resonance varies. In the cases of spin-orbit resonance, gravitationally produced tides cause the resonances. In cases of orbital resonance, the gravitational forces from two bodies combine to produce either resonant gaps or stable, coorbital points where small particles accumulate. Examples of all these types of resonance can be found within the solar system.

The most familiar example of spin-orbit resonance is the motion of the Earth's only natural satellite, the Moon. Tidal stresses on the Moon from Earth have locked the Moon in its spin-orbit resonance so that only one side faces Earth. Earth dwellers are inclined to think that the Moon does not spin or rotate, but if viewed from far out in space, the Moon would be seen to spin once for every orbit it makes of the Earth. As an illustration, the Moon has phases because sunlight reaches all points of the Moon. This indicates that, as viewed from the Sun, the Moon spins once a month, which is exactly the same time that it takes to orbit the Earth, and is therefore in a 1:1, or synchronous spin-orbit, resonance with Earth. The Moon is often said to be "tidally locked" to the Earth because of this resonance and its cause.

The Moon is not the only secondary satellite in the solar system to exhibit synchronous rotation. Tidal locking appears to be the rule for all satellites close to a planet. In fact, Phobos and Deimos are in synchronous rotation around Mars. The four Galilean moons of Jupiter, which are some of the largest satellites in the solar system, also exhibit a 1:1 spin-orbit resonance. Of the many other satellites of Jupiter, two others that had their rotational periods measured; one of those, the closest one to Jupiter, is synchronous. A similar situation exists for the satellites of Saturn, where eight are known to have synchronous rotations; eight other sizable satellites have not been measured. At least two Saturnian satellites are nonsynchronous. The largest satellites of Uranus and Neptune are also in synchronous rotation, and the smaller ones have not yet been measured. Pluto's companion Charon not only is in synchronous rotation around Pluto but also is large and close enough to have caused Pluto's rotation to be synchronous with Charon's orbit. Thus, all planets with satellites have examples of synchronously rotating moons.

Mercury, although lacking any moon, also exhibits spin-orbit resonance. Mercury spins three times for every two orbits it makes of the Sun. This 3:2 spin-orbit resonance is related to the unusually elongated, elliptical orbit of Mercury. As a result of its resonance condition, whenever Mercury is at perihelion, the same point on the planet is either facing directly toward or away from the Sun. The Mariner 10 spacecraft identified a huge feature, the Caloris Basin, at this point on the surface and some strange surface features, dubbed Weird Terrain, on the planet's opposite side. These discoveries suggest that a huge, ancient impact that nearly tore the planet apart made one side of the planet heavier than the other and probably elongated and tilted the orbit. Tidal effects over the years have slowed the rotation of the planet so that whenever Mercury is at perihelion, its heavy side points either toward or away from the Sun, as a tidal bulge would. The fact that the spin is not synchronous with the orbit, as it is for Earth's moon, is most likely the result of Mercury's large mass and elongated orbit, which brings it considerably closer to the Sun at perihelion than at aphelion.

Evidence of orbital resonance was discovered in the asteroid belt between Mars and Jupiter in 1866 by Daniel Kirkwood. As he studied the orbits of the asteroids, Kirkwood discovered gaps in an otherwise congested region of space. Since Kirkwood's original discovery of gaps at 2:1, 3:1, and 4:1 resonances with Jupiter, at least five other gaps have been identified. It is apparent that the strong and repeated pull of Jupiter destabilized orbits of asteroids with these periods and opened up the Kirkwood gaps.

Divisions in Saturn's rings have a cause similar to the Kirkwood gaps. Gian Domenico Cassini first observed the largest gap, the Cassini division, in 1675; in 1867, Kirkwood discovered that the Cassini division has a 2:1 orbital resonance with the satellite Mimas. Kirkwood also showed that the Cassini division was in a 3:1 resonance with the satellite Enceladus, a 4:1 resonance with the moon Tethys, and a 6:1 resonance with the satellite Dione, although the Mimas resonance is probably more significant because that coupling is stronger and more frequent. In addition to the Cassini division, there are gaps in the A ring at resonances with the satellites Janus (S10) and Epimetheus (S11). Moreover, the edges of the A and B rings, which are very well defined, occur at resonance locations.

There are still some mysteries to be found in the asteroid belt and Saturn's rings. Surprisingly, at the Jupiter 3:2 resonance location in the asteroid belt, there is accumulated material instead of an empty gap. In the Cassini division, there are ringlets, which may be spiral density waves excited by resonances with the satellite Iapetus. Other details of the structure and shape of the divisions are still not well understood and may be aspects of density waves and chaotic behavior. Nevertheless, these are details, and the main features must be caused by the simple resonances.

Jupiter and Saturn also have several examples of coorbital satellites. In Jupiter's orbit around the Sun, there are clumps of asteroids one-sixth of an orbit ahead and behind Jupiter at the Lagrangian points. These coorbital asteroids are called Trojan asteroids. In the Saturnian system, the satellite Tethys has two Lagrangian coorbital satellites, Telesto (following Tethys) and Calypso (leading Tethys). In addition, the satellite Dione has a coorbital satellite named Helene at the leading Lagrangian point. Other examples are expected to exist, but they have not yet been observed.

Two of Saturn's satellites, Janus and Epimetheus, are also coorbital, but in a different way from that of the Lagrangian coorbitals. Janus and Epimetheus have orbits that are so close together that their gravitational attraction for each other is sufficient for them to interchange orbits without colliding. The difference between this case and that of the Lagrangian coorbital satellites results from the fact that Janus and Epimetheus are nearly the same size and that one satellite does not always lead the other.

CONTEXT

The result of the Moon's spin-orbit resonance has been known ever since humans became aware of the world around them, but the mechanism for understanding why such a resonance would occur was not discovered until Sir Isaac Newton formulated his laws. In fact, tides were explained in his *Philosophiae Naturalis Principia Mathematica* (1687; *Newton's Principia: The Mathematical Principals of Natural Philosophy*, 1846), in which he first used his laws publicly to explain Johannes Kepler's laws of planetary motion. It was thought that most satellites would have similar resonances with their planets, but this extrapolation needed to be confirmed. Planetary probes have visited all planets except Pluto, and the New Horizons spacecraft is on its way to a flyby of the Pluto-Charon system in 2015. With a few notable exceptions, many satellites have exhibited spin-orbit resonance. Unfortunately, all the satellites were not able to be studied thoroughly to determine their spin rates; as a result, a complete understanding of which moons are tidally locked to their planets, and how, will not be within reach until additional probes or improved technology becomes available. The Galileo probe improved our understanding of Jupiter's satellites, and Cassini added to what is known about Saturn's.

Interestingly, Mercury generally was thought to be tidally locked to the Sun ever since Giovanni Schiaparelli made crude maps of the surface in the 1880's. This opinion seemed to be confirmed by later Earth-based observations that were carried out prior to the early 1960's. It was not until Doppler radar techniques were applied to Mercury in 1965 that the 3:2 spin-orbit resonance was discovered. In this study, radar signals were sent from the 300-meter Arecibo radio telescope in Puerto Rico and bounced off Mercury. The change in the signal's frequency (the Doppler effect) proved that Mercury rotates once in 58.65 Earth days instead of the 88 days that it takes to orbit the Sun. It was not until the three Mariner 10 flybys of Mercury in 1974 and early 1975 that the Caloris Basin and the Weird Terrain were discovered and the orbital resonance was confirmed.

Discovery of the Cassini division in 1675 also predates Newton's laws, which are essential to explain the division. It was not until Kirkwood's discovery of resonance conditions in 1866 and 1867 that a reasonable explanation for formation of this division was offered. Kirkwood's model for the Cassini division and gaps in the asteroid belt was thought to be adequate until the Voyager data became available in 1981. Images of the Cassini division from the two Voyager spacecraft revealed a number of unexpected details. These details require a more sophisticated application of Newtonian mechanics and provide a testing ground for density wave theories and theories of chaotic behavior. These theories could help explain the structure of galaxies.

In contrast to previous cases, Lagrange predicted in 1772 the location of stable, coorbital companions to Jupiter and other planets. In 1906, Jupiter's coorbital companions were found. The search continues with nearly two hundred such asteroids discovered to date and perhaps ten times that number orbiting at Jupiter's Lagrangian points. The coorbital satellites in Saturn's system were undiscovered until the Voyager flybys in 1980 and 1981. In fact, Lagrangian coorbital satellites may exist for Earth, Mars, and other large bodies.

Larry M. Browning

FURTHER READING

Elliot, James, and Richard Kerr. *Rings: Discoveries from Galileo to Voyager*. Cambridge, Mass.: MIT Press, 1984. A detailed discussion of the structure and formation of planetary rings. Discusses the connection between the Cassini division, spiral density waves, and galaxies.

Fowles, Grant R., and George L. Cassiday. *Analytic Mechanics*. 7th ed. New York: Brooks/Cole, 2004. A college textbook for a second course in Newtonian mechanics. Particularly strong on theory and applications of orbital motion. Requires knowledge of advanced calculus.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find the standard treatise for typical ASTRO 101 classes too low level. Suitable for an audience with varied science and mathematical backgrounds. Covers all topics from solar-system objects to cosmology.

Miner, Ellis D., Randii R. Wessen, and Jeffrey N. Cuzzi. *Planetary Ring Systems*. New York: Springer Praxis, 2006. Perhaps the most comprehensive text on planetary ring systems that is also accessible to the general reader. Provides interpretation of Pioneer, Voyager, Galileo, and Cassini data and observations. Extensive notes, tables, figures, and references. Accessible to a nonexpert, scientifically inclined audience.

Shu, Frank H. *The Physical Universe: An Introduction to Astronomy*. Mill Valley, Calif.: University Science Books, 1982. Although somewhat dated and more mathematical than the other references, this book contains an extensive and enlightening discussion of resonances in the solar system. Also noteworthy is the discussion of spiral density waves and their application to the Cassini division by one of the pioneers of that theory.

Time-Life Books. *Comets, Asteroids, and Meteorites*. Alexandria, Va.: Author, 1990. This volume discusses the Trojan asteroids orbiting with Jupiter and the theories about the formation of the Kirkwood gaps. Very readable and well illustrated.

_______. *The Far Planets*. Alexandria, Va.: Author, 1988. Notable for its pictures and informative illustrations of Saturn's ring system. Published before the August, 1989, flyby of Neptune by Voyager 2.

_______. *The Near Planets*. Alexandria, Va.: Author, 1989. This volume provides a very readable summary of the exploration of Mercury and its spin-orbit resonance. A good source of scientifically accurate illustrations.

Wagner, Jeffrey K. *Introduction to the Solar System*. Philadelphia: Saunders College Publishing, 1991. A well-written and up-to-date discussion of all aspects of the solar system, especially spin-orbit coupling.

See also: Planetary Orbits; Earth's Rotation; Planetary Rotation.

Planetary Ring Systems

Category: Planets and Planetology

A planetary ring system consists of enormous numbers of relatively small particles that fan out from a planet in the form of a disk, orbiting as a complex unit around that planet. Planetary rings are common in the outer solar system; each of the four "gas giant" planets has a ring system of different age and degree of complexity.

OVERVIEW

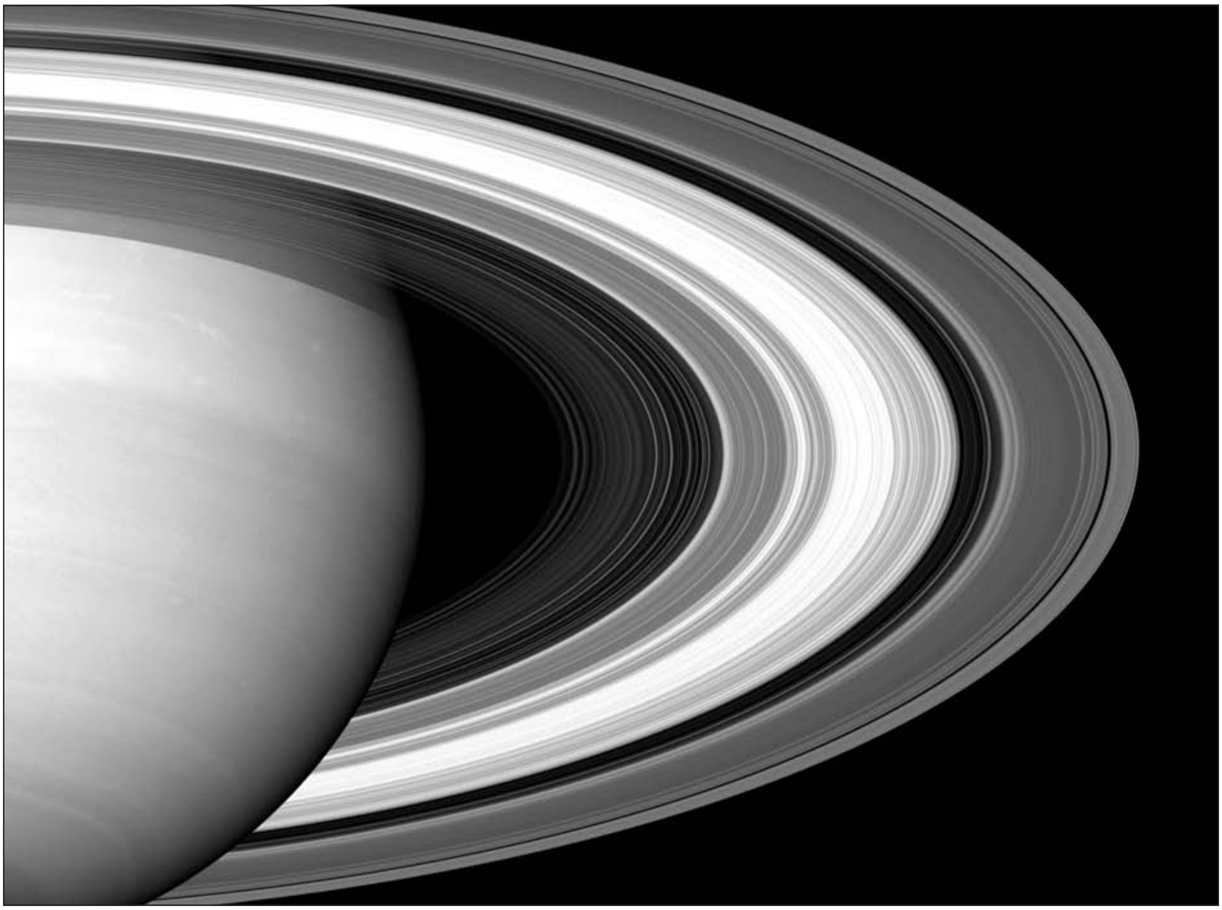
The rings of Saturn are stunning, perhaps among the most beautiful features to observe in the night sky. They were first discovered in 1610 by Galileo in Padua, Italy. Galileo's telescope was not the best of astronomical instruments. What he sketched were two spheres, one on each side of the planet. Galileo thought he had discovered that Saturn was a triple planet. The matter was clarified in 1655, when Dutch observer and telescope maker Christiaan Huygens clearly saw Saturn's rings through his improved instrument. Later, in 1675, Gian Domenico Cassini, supervisor of the Paris Observatory, discovered that there was structure to the rings. An opening in the ring exists about two-thirds of the way out from the planet. This gap became known as the Cassini division. The outer ring was called the A ring and the inner ring, the B ring.

The rings of Saturn were of intense interest to scientists throughout the eighteenth and nineteenth centuries. New gaps and subrings were subsequently discovered and named. In 1789, English scientist Sir William Herschel estimated the rings to be no more than 500 kilometers thick. By 1850, astronomers' telescopes could resolve that the rings were largely transparent and that the edge of the planet could be seen through the formation. In 1848, French mathematician Édouard Roche rigorously proved that if a satellite orbited too close to a planet, tidal forces from that planet would tear it apart into small pieces. The planet also would not allow a satellite to form from small pieces inside this distance. This limit is known as the "Roche limit." The rings of Saturn fell inside the Roche limit. Roche boldly suggested that Saturn had captured a small satellite that had then been broken into billions of tiny pieces by the gas giant's gravitational forces.

In 1857, the University of Cambridge offered a prize to settle the question of whether the rings were rigid, fluid, or made up of small pieces of matter "not mutually coherent." Scottish physicist James Clerk Maxwell presented a mathematical argument that won the prize. In his proof, Maxwell demonstrated that any solid ring would be torn apart by gravitational forces. He also demonstrated that the rings could not be a liquid. Thus, the rings of Saturn were composed of countless individual particles, each in its own independent orbit around the planet.

From the mid-nineteenth century until 1979, discoveries about Saturn's ring system were largely limited to finding new divisions, and it was postulated that the rings were no more than 15 kilometers thick. In 1977, rings around the planet Uranus were observed using a stellar occultation method. In 1979, the National Aeronautics and Space Administration's (NASA's) twin Voyager spacecraft discovered rings around Jupiter. Voyager 2 confirmed Uranus's rings and, in 1989, discovered diminutive ring arcs around Neptune.

Much about all the planetary ring systems can be learned from the study of Saturn's rings, although the composition of each of the giant planets' rings has not been determined as completely as has Saturn's. The Saturnian ring system is composed almost completely of water ice fragments ranging microscopic particles to boulder-sized bodies, whereas the composition of the rings of Jupiter, Uranus, and Neptune is not fully known. Saturn's rings begin only 32,000 kilometers above the planet's clouds and extend outward to 230,000 kilometers. The rings are no more than a few hundred meters thick. There are seven major rings separated by what are still known as "gaps" (including Cassini's gap between the A and B rings). A closer inspection by Voyager revealed that the major rings consist of tens of thousands of ringlets that resemble grooves on a vinyl phonograph record. The ringlets are not cleanly separated from one another but appear to exhibit the property of wave propagation through the structure. There are spiral density waves as well as bending waves. Each ringlet differs in width from 100 kilometers to less than 1 kilometer.



Saturn's ring system was once thought to be the only one in the solar system; it is now known to be only the most pronounced of the ring systems of the four Jovian planets. (NASA/JPL)

The rings are composed almost exclusively of ice fragments whose diameters range from submicron-size to 10 meters or more. Denser regions of the rings appear to be composed largely of smaller particles, while in the gaps, the larger, meter-sized particles dominate. There may be many ring particles as large as 50 kilometers across within the ring system, although the extent of these particles is not known. As of 2009, the Cassini spacecraft continued to provide images from its orbit about Saturn that reveal more structure to the planet's complex ring system.

One of the most astonishing discoveries made by the Voyagers was the tremendous dynamic complexity of the ring system. An unexpected dynamic effect found by Voyager was the spokelike effect seen in movies assembled out of individual sequences of images taken by the spacecraft. Radial spokes extend outward like spokes on a bicycle wheel. They rotate with the rings. Spokes tended to form, widen, then disappear after about an hour. The most widely accepted theory is that spokes are formed by electrostatic forces between the submicron-size particles and the planet. They are short-lived because the orbital period of the inner particles is faster than that of the outer particles of the spoke.

Another of the Voyager discoveries was the existence of what become known as "shepherd satellites." These are tiny satellites (not readily visible from Earth-based telescopes) that orbit on the outside fringes

Gian Domenico Cassini, Planetary Explorer

In July, 1664, a professor of astronomy at the University of Bologna, Gian Domenico Cassini, made his first major observation: Jupiter was not a perfect sphere but instead was flattened at its poles. Over the next few years he measured Jupiter's rotational period, observed its moons, and discovered discrepancies in his own measurements that, at first, he attributed to light having a finite speed. (However, he later appears to have rejected his own idea, and in 1676 Danish astronomer Ole Rømer would use Cassini's measurements to calculate the speed of light.)

In 1666, Cassini observed surface features on Mars, including Syrtis Major. Again, he measured the planet's rotational period using these features and produced a value of 24 hours and 40 minutes—within 3 minutes of the period now accepted. He also attempted to determine the rotational period of Venus, which he calculated as 23 hours, 20 minutes. What Cassini observed to produce this conclusion is unclear; Venus is entirely covered by bright clouds, and its rotational period (about 243 days) became clear only with the advent of radar.


(Library of Congress)

By this time Cassini's measurements had made him famous throughout Europe, and he came to the attention of Jean-Bapiste Colbert, the French minister of finance. At Colbert's suggestion, King Louis XIV invited Cassini to head the new Paris observatory in 1669. In Paris, Cassini discovered two moons of Saturn, Iapetus and Rhea. In 1675, he recognized that Saturn's ring was divided, separated by a dark gap now called the Cassini division. In 1677, Cassini demonstrated that Saturn was flattened at its poles, and in 1684 he discovered two more moons, Dione and Thetys. In 1705, he correctly suggested that Saturn's ring might not be a solid disk but rather a swarm of small objects orbiting the planet. Cassini also observed several comets between 1672 and 1707, as well as Jupiter's Great Red Spot.

During his years at the Paris observatory, Cassini organized a renowned group of astronomers—including Christiaan Huygens, Ole Rømer, and others—known as the Paris School. Trained in engineering, he published several works on flood control, served as inspector of water and waterways, and became superintendent of the Fort Urban fortifications. By 1711 he was blind, and he died in Paris on September 14, 1712. His son Jacques (1677-1756) became head of the Paris observatory, and his grandson César-François Cassini de Thury (1714-1784) and great-grandson Jacques-Dominique de Cassini (1748-1845) also became noted astronomers.

of the rings within some of the gaps. The minuscule gravitational field of these tiny satellites is enough to push ring particles back into the rings. They help define the rings' outer edges (hence the name "shepherd"). Shepherding satellites also exhibit some rather interesting effects on the rings. In the case of the F ring, two shepherding satellites (1980S26 and 1980S27, later named Pandora and Prometheus, respectively) confine the ring. The F ring appears to be discontinuous in places, is intertwined in others, and is knotted and lumpy in others. The current theory is that the F ring's complexity is caused by the slight eccentricity in the orbits of the two shepherding satellites. It is speculated that over time, variant gravitational interactions cause the structural convolutions. Neptune's ring system of partial arcs exhibited similar discontinuities and irregularities, but there was not enough resolution or coverage to discern shepherding satellites.

Even before the Voyager encounter and the discovery of the superior structural details, a concept known as satellite "ring resonance" was advanced to explain what could be observed from Earth-based telescopes. This concept indicates that the shape of the rings and the location of the gaps are determined not only by the orbits of individual ring particles but also by gravitational influences of Saturn and its major satellites. This theory postulates that Mimas and Enceladus have a particular influence on particles of the rings so that the Cassini division is created by the gravitational interaction of the satellites on the particles in the ring. The discoveries of Voyager supported part of the resonance theory, but shepherding satellites and other gaps that had no resonance explanation in Saturn's system and that of the other giant planets left some questions about its ultimate effect. Those questions were left for the Cassini orbiter to investigate.

The rings of Saturn are flattened along the equatorial belt because of countless energy-dissipating collisions between particles that have, over the millennia, all but eliminated vertical displacements. Such collisions do not affect the circular orbital motion of the particles; hence, the net effect is a disklike flattening.

The ring system at Jupiter was found to be faint and largely made of dust. Between Voyager and Galileo spacecraft observations, it has been determined that Jupiter's ring system has four components. There is an inner Halo Ring, a relatively thick torus-shaped collection of particles. The Main Ring is very thin and relatively bright. Then there are two Gossamer Rings that are rather wide, thick, and very faint. One is associated with the satellite Amalthea, and the other with the satellite Thebe; these Gossamer Rings are believed to be composed of material coming from their associated satellites.

Astronomers had not expected Jupiter to have rings. Before Voyager, rings there had never been observed from Earth. As Voyager passed behind Jupiter on its outbound journey, it looked back on the giant planet and photographed the ring particles in reflected sunlight. It found that the Jovian rings absorbed all but one ten-thousandth of the sunlight incident upon them. There are two main rings around Jupiter, at 47,000 and 53,000 kilometers above the cloud tops. One ring is 5,000 kilometers wide, while the outer ring is only 800 kilometers in width. There is a torus of thick particles that is not particularly bright, and the small satellites Amalthea and Thebe produce the particles that compose the two Gossamer Rings. Voyager was unable to detect ring thicknesses precisely, but they were estimated to be between 1 and 30 kilometers wide. Voyager scientists estimated that most of the Jovian ring particles are probably made up of dust-size pieces, each in an individual orbit around Jupiter with an orbital period from five to seven hours. Because the tiny, dark particles are in unstable orbits and are constantly falling in toward Jupiter, the rings are probably constantly renewed by the tiny satellite Adrastea, which was also discovered by Voyager.

The rings of Uranus were actually detected before the Voyager spacecraft arrived at that planet. Voyager discovered five well-defined rings around Uranus and four other, less defined rings. The innermost Uranian ring appeared quite compact and dense, while the outermost ring was quite diffuse. Aside from the nine distinct rings around Uranus, from fifty to one hundred nebulous dust bands were discovered in outer orbits containing very small particles.

It is speculated that the Uranian ring system began when ten to twelve small satellites less than 200 kilometers in diameter began to break up. As they broke up in low orbit, the fragments began to collide with one another, forming the dust bands and main ring system. Thus, the Uranian ring system may be constantly replenished so that it may have a lifetime of millions of years.

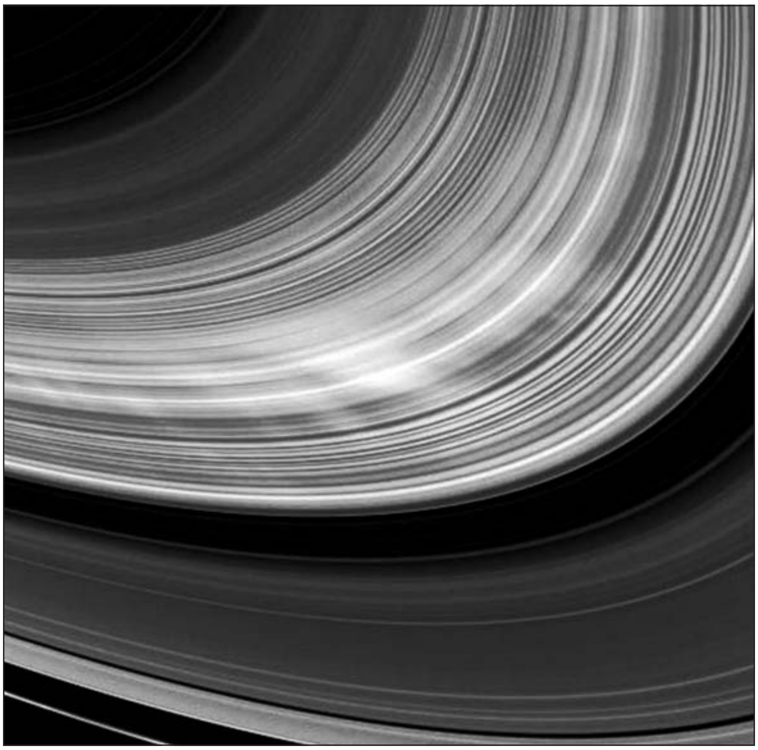
Voyager 2 flew past Neptune in August, 1989, and discovered that this gas giant planet also has a distinct ring system. Voyager found five incomplete ring arcs around Neptune. The exceptionally thin, bright, innermost ring of Neptune was found to be denser than the others, some 17 kilometers wide with a distinct, more diffuse component extending 50 kilometers in toward Neptune. In the outermost rings, distinct points of light were observed, suggesting that the rings are embedded with tiny, icy moonlets, which may act as shepherding agents, as in the Saturnian system. These tiny moonlets are estimated to be about 10 or 15 kilometers in diameter. The outer rings of Neptune are much wider than the inner rings; the third ring is quite dispersed and tenuous but is some 2,500 kilometers across.

The rings of all the giant planets may have been formed by one of several mechanisms. Ring particles may have been accumulated from the breakup of a satellite destroyed by tidal forces; they may have been created when a satellite and asteroid or comet collided in orbit; or the rings may be merely particles left over from the formation of the planet inside the Roche limit that could not accrete into a satellite because of the tidal forces. Why the rings of Saturn are so different from those of the other giant planets may be explained by different material origins. It is possible that the rings around Saturn were created when a satellite of icy origin broke up, while the dark rings of Jupiter represent the remains of a body made of much darker material. None of these conjectures will be proved until pieces of the rings can be directly analyzed.

METHODS OF STUDY

The rings around Saturn are the best understood of the planetary ring systems because their existence was well known long before the Voyagers were launched. Much of the information gained at Saturn by the Voyagers and later the Cassini orbiter can be applied to the other giant planet ring systems, even though the other planets' rings appear much different.

As an orbital platform capable of examining the complex ring system around Saturn with sensors viewing in several portions of the electromagnetic spectrum in addition to the visible, the Cassini spacecraft verified a number of the Voyager findings and provided even newer insights into ring dynamics, the spokes phenomenon, and gravitational interactions between ring particles and shepherding satellites. Cas-



Saturn's B ring, as seen from the Cassini spacecraft in 2008. (NASA/JPL/Space Science Institute) sini arrived in Saturn orbit on July 1, 2004, and finished a four-year primary mission. Fortunately, Cassini was given an extended mission in light of the spacecraft's good health and constant stream of new and exciting data as well as ongoing stunning photography. In addition to all the refinements of Saturn's ring structure, Cassini produced hints of a tenuous ring system around Saturn's second-largest satellite, Rhea. Cassini used its Magnetospheric Imaging Instrument (MIMI) to infer these thin rings based on a depletion of energetic electrons near that satellite. MIMI data indicated three drops in electron concentration symmetrically located about Rhea, suggesting that particles from decimeters to meters in size were absorbing the electrons. The rings were too tenuous and dark to image directly.

The evolution of planetary ring systems may all be the same. A careful analysis of ring system material should eventually demonstrate this fact. The dynamics of ring systems seem to be relatively similar. Braided and discontinuous rings were seen in all ring systems. Shepherding satellites have been discovered in at least two of the systems. Spokes seemed to be confined to Saturn, which may have something to do with the fact that the mass of material in the Saturnian system was significantly higher than in any of the others.

The question of a ring system's total mass is an important one and awaits resolution. It may be that ring systems have definite lives. If there is no shepherding action, many of the ring particles may be spun out of the ring system by multiple piece encounters, along with random gravitational and tidal effects from the planet and its larger satellites. It may be that the complex Saturnian ring system is relatively young when compared with the other planets' rings; it appears to be better developed. A robust, more massive ring system, thus, may be a sign only of its youth, while a devitalized ring system such as seen on the other gas giant planets may be evidence of their advanced age. This hypothesis, however, remains speculative at best.

A study of ring systems has direct applications to the study of newly forming planetary systems. In this comparison, Saturn can be visualized as a forming star, while the ring particles can be seen as the developing stellar nebula of dust and particles. Such a system has been theorized for our solar system's development some 4.5 billion years ago. In this comparison, the behavior and dynamics of the particles in planetary ring systems can be compared with the early solar system. Even though the Saturnian ring particles are prevented from accretion by confinement within the Roche limit, some comparisons may be made with other dynamic system components.

Such a comparison is not only valuable for direct association to the solar system but also valuable elsewhere in the galaxy. A careful study of the dynamics of planetary ring systems may ultimately be used to calculate the probability of other nebular systems developing around other solar systems. An extension of such estimates will make possible more accurate calculation of the number of planetary systems, where extrasolar planets develop with respect to their star, and perhaps even how many planets could be expected to develop.

CONTEXT

The Saturnian rings have long held a special place in the history of science. Because of their spectacular beauty, they have always been considered to be the crown jewel of the solar system. The application of science to the study of those rings began with sketches by the first scientist to view them, the reputed inventor of the telescope, Galileo. The rings became the focal point for one of the first evaluations of the nature of the solar system from Earth. Mathematical evaluation of the rings accomplished by Maxwell foretold of an era of scientific investigation without direct visitation.

The Pioneer, Voyager, Galileo, and Cassini spacecraft extended humanity's reach to the ringed planets and greatly enhanced the methods of remote study. Evaluating the mass of data returned by robotic probes will continue to keep theorists busy for years to come. Enigmatic spokes, braided rings, and discontinuities each present problems that are still far from being solved. Ring resonance and the concept of ringlets as manifestations of a kind of particulate periodicity also have recently been modeled. Some of remaining dynamic and compositional questions will almost certainly require direct return of samples and other visits by robotic probes. Explorations of the rings will undoubtedly continue, and will provide astronomers and planetary scientists with exciting scientific opportunities for decades.

Dennis Chamberland

FURTHER READING

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. A popular work on solar system astronomy and planetary exploration, filled with color diagrams and photographs. Covers discoveries through the Mars Pathfinder and Galileo missions. Provokes excitement in the general reader, who gains an explanation of the need for greater understanding of the universe.

Bortolotti, Dan. *Exploring Saturn*. New York: Firefly Books, 2003. Full of charts, photographs, a section on observing Saturn, and a historical development of an understanding of the Saturn system from antiquity to the launch of Cassini. For younger readers.

Bredeson, Carmen. *NASA Planetary Spacecraft: Galileo, Magellan, Pathfinder, and Voyager*. New York: Enslow, 2000. This book is part of Enslow's Countdown to Space series and provides an overview of NASA planetary exploration during the last two decades of the twentieth century. Designed for younger readers but suitable for all audiences.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that had investigated all planets with the exception of Pluto by the publication date. Takes an astrophysical approach to place our solar system in a broader context as just one member of similar systems throughout the universe.

Harland, David M. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002. A technical description of the Cassini program, including its science goals and the instruments used to accomplish those goals. Written before Cassini arrived at Saturn. Provides a historical review of pre-Cassini knowledge of the Saturn system.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. The material on Jupiter, Saturn, Uranus, and Neptune describes all four ring systems and discusses spacecraft exploration of them. Takes a comparative planetology approach rather than presenting individual chapters on each major object in the solar system.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Accessible to the serious general audience.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find the standard text for introductory astronomy courses too basic. Covers all topics from solar-system objects to cosmology. Suitable for an audience with varied science and mathematical backgrounds.

Miner, Ellis D., Randii R. Wessen, and Jeffrey N. Cuzzi. *Planetary Ring Systems*. New York: Springer Praxis, 2006. Perhaps the most comprehensive text on planetary ring systems that is also accessible to the scientifically inclined reader. Provides interpretation of Pioneer, Voyager, Galileo, and Cassini data and observations. Extensive notes, tables, figures, and references.

Morrison, David. *Voyages to Saturn*. NASA SP-451. Washington, D.C.: Government Printing Office, 1982. This easy-to-read classic NASA publication examines the Voyager encounters with Saturn. Includes details of Voyager's Jupiter encounter and the Jupiter rings. Outlines the history of rings and ring theory. Includes photographs and an index.

See also: Jupiter's Ring System; Neptune's Ring System; Saturn's Ring System; Uranus's Rings.

Planetary Rotation

Category: Planets and Planetology

The rate and direction of a planet's rotation record major incidents from that planet's past and determine, to a large extent, the climate that a planet will experience. Study of the rotation also yields clues to the generation of the planet's magnetic field.

OVERVIEW

Physicists recognize two possible changes in an extended rigid object. Such an object can change position or can change its orientation (or both). The first is called displacement and the second, rotation. A total description of the motion of an object requires a determination of the movement of that extended body's center of mass and the specification of three angles relative to a coordinate system fixed with its origin at the center of mass within the body of the rigid object. That is to say that full space motion of an object requires specification of six degrees of freedom (three coordinates of movement and three angles defining orientation).

Much of what is known about planetary rotation has been learned, not surprisingly, by observation and study of our own planet, Earth. Earth rotates on its axis, but the time to spin once about its axis depends on the point of reference used. If stars are used as an example, Earth rotates twenty-three hours, fifty-six minutes, and four seconds back to alignment with the fixed stars. This rotation is known as the sidereal period. If, on the other hand, the Sun is used as the reference, the Earth rotates in twenty-four hours, from one noon to the next; this rotation is a solar day. The reason that the solar day and the sidereal period are not the same is that Earth revolves around the Sun. It takes a shorter period of time for the Earth to realign with the stars than it does for it to realign with the Sun.

The Earth's axis wobbles. If the positions of stars were observed overhead at the North Pole for several months, one would notice that they change position slightly. Actually, they are not changing position; rather, the spin axis of Earth is changing. It is not known precisely what causes this "nutation," but winds and seismic activity may be factors.

The tides are produced by gravitational attraction of Earth's oceans by the Sun and the Moon. This attraction also produces tides in the atmosphere and the solid Earth, but these two tides are not normally noticed. As ocean water moves across shallow seas on the Earth, it rubs against the bottom of the basin. This action produces a friction that slows down the Earth, but angular momentum must be conserved. Angular momentum is the product of an object's mass, angular velocity, and an appropriate measure of its size. For rotation, the appropriate measure would be the Earth's radius. However, for revolution that measure would be its orbital distance from the Sun. The Moon is locked into its present rotation and cannot speed up or slow down. Since the Moon and Earth are gravitationally bound to each other, that is where the momentum is conserved. The Earth loses angular momentum by slowing down, but the Earth-Moon system gains it. For this to occur, the Earth-Moon distance is increasing by roughly 4 centimeters per year. At one time, the Moon rotated much faster, but tides in the solid Moon produced by Earth's gravitational pull slowed the Moon to its present state with the same face always pointed toward Earth.

Earth has been slowing down for hundreds of millions of years. Scientists look at growth rings in coral fossils 400 million years old; each ring is deposited in a day and varies in thickness depending on the time of year. Four hundred rings were found for one year's growth, which corresponds to a twenty-two-hour day. Presumably, Earth was rotating even faster before that time. Observations of other objects in the solar system indicate that a rotation of six to ten hours is normal. Any rate less than that is caused by a process, such as the tides, slowing down the object.

Earth's axis of rotation is tipped roughly 23.5° from the perpendicular to Earth's orbital plane. This tip produces the seasons. When Earth's Northern Hemisphere is pointing toward the Sun, it is summer, while the Southern Hemisphere endures winter. As Earth orbits the Sun, six months later, the Southern Hemisphere points toward the Sun. It will experience summer, and the Northern Hemisphere will undergo winter. If Earth were not tilted, there would be no drastic changes in the weather.

Rotation of Earth also creates the Coriolis effect. If a freely suspended pendulum is observed over several hours, it appears to shift its position relative to the ground. In fact, it is not the pendulum but Earth that is rotating beneath the pendulum. This effect also is important in directing ocean currents. In the Northern Hemisphere, as the current moves north, the Coriolis effect pushes water to the right or eastward. Water moving east is pushed south, water moving south is pushed west, and water moving west is moved to the north. This produces clockwise motion of major ocean currents in the Northern Hemisphere. In the Southern Hemisphere, the motion is counterclockwise. The effect also explains the way in which areas of atmospheric high and low pressure spiral clockwise and counterclockwise, respectively, in the Northern Hemisphere. (Water spiraling down a drain does not do so because of the Coriolis effect, however; the Coriolis force is much too weak in comparison to other forces affecting the water's motion.)

Several millennia ago, astronomers of the Middle East discovered that Earth precesses. Precession is a slow change in the direction of a planet's axis of rotation. In general, precession is a change in the direction of the axis of rotation of an object resulting from the application of a torque; torque is any force not acting through a body's center of mass. If a toy top is spun, for example, its spin axis will point in one direction, but as the top slows down, the spin axis moves in a circular motion. The top takes several seconds to precess through one cycle. Earth's period of precession is much greater, roughly twenty-six thousand years. The reason that the top precesses results from the fact that gravity is trying to pull it over, but the top responds by precessing in a direction perpendicular to the spin axis and the direction of the force. Earth precesses because the Sun and the Moon are gravitationally pulling on Earth's equatorial bulge. This pull tends to change the direction of Earth's rotation, which responds by precessing.

If one imagines that the heavens are actually a sphere—the celestial sphere—with the stars and other objects on it, then one can project Earth's spin axis onto the celestial sphere. It passes near the pole star, Polaris, in the constellation Ursa Minor. In the past, because of Earth's precession, other stars such as Vega in the constellation Lyra have been the pole star. In the future, as Earth continues to precess, these stars will again become Earth's pole star.

As a result of sending spacecraft to other planets, some surprising information has been gained, especially about their rotation. Before 1965, the rotation rate of Mercury was not known. Astronomers assumed that it was the same as its rate of revolution—in other words, that it kept the same face toward the Sun. Mariner 10 determined that Mercury's period of rotation is 58.65 days. Mercury's period of revolution is 87.97 days. This means that for every two revolutions that Mercury makes around the Sun, it rotates about its axis three times. Mercury is locked into what is called a 2:3 spin-orbit coupling. Earth's moon is locked into a 1:1 coupling. It would take tremendous energy to make either of them rotate at a different rate.

In 1974, Mariner 10 arrived at Mercury to take photographs and to measure other characteristics. It was suggested that, although the spacecraft could not be put into orbit around Mercury, it could be placed in an orbit with a period that returned it to Mercury once every two revolutions of the planet. (This calculation was made prior to the discovery of Mercury's period of rotation.) As Mariner 10 approached, it took pictures of one face. Two revolutions later, it returned and photographed the same face, and again two revolutions later, the same face was photographed. Photographs were taken of only one hemisphere of Mercury.

The rotation of Venus was another mystery, since the planet's surface could not be observed because of its dense cloud cover. That cloud cover is also uniform in color, which prevents determination of the rotation by looking at a feature, such as a spot. Radar pulses sent to Venus and reflected back to antennas on Earth have been used to determine the rate. When the radio pulse (electromagnetic radiation) is transmitted, it has a specific frequency. As the beam strikes the surface of the planet, it is reflected,

Planet Shapes and Rotational Periods and Velocities

Planet	*Oblateness*	*Rotational Period*	*Surface Velocity of Rotation*
Mercury	0.0	59 days	10.8
Venus	0.0	–243 days	6.5
Earth	0.003	24 hours	1,670.0
Mars	0.009	24.5 hours	866.0
Jupiter	0.06	10 hours	45,087.0
Saturn	0.1	10.5 hours	36,887.0
Uranus	0.06	–17 hours	14,794.0
Neptune	0.02	16 hours	9,794.0
Pluto (dwarf)	?	–6.5 days	128.0

Notes: Oblateness, or ellipticity, is a measure of the planet's flatness, or deviation from a perfect sphere. Rotational periods are rounded to the nearest half day or half hour; a minus sign signifies retrograde rotation.

but there is a shift in the frequency depending on the velocity of the surface at the reflection point. This shift is a result of the Doppler effect. For example, if a siren is heard approaching someone, its pitch rises; as it passes, the pitch becomes lower. In fact, the sound waves emitted by the siren have not changed, but they have been compressed relative to the person hearing them as the siren approaches, and they are perceived as expanded, or stretched apart (hence, lower in pitch), as the siren recedes from the hearer. A similar process affects the radar pulse. If the surface is moving toward Earth, the frequency of the reflected beam is raised in proportion to the surface velocity. As it is moves away, the frequency drops in proportion to velocity. The received pulse is analyzed for frequency shift, which is converted into a rotation rate.

Since Venus is similar to Earth in size and mass, scientists assumed that its rotation rate would be similar to that of Earth. They soon realized, however, that Venus rotates once every 243 days (sidereal), which is greater than its orbital period. They found that Earth and most other planets rotate in the counterclockwise direction, as viewed from the North Pole of the Earth, whereas Venus rotates in a clockwise direction; Venus's rotation is retrograde. Astronomers have studied other objects in the solar system to develop models for the origin of Venus's retrograde rotation. One possible explanation for this retrograde rotation is that a large object, perhaps the size of Mars, struck a glancing blow to Venus. This impact acted like a brake to slow Venus down and reverse its direction. (A similar event is thought to have occurred to the Earth, but in this case, material was thrown out from the Earth and some of that material was later pulled together by gravity to form the Moon.)

Mars rotates once every twenty-four hours and 37.5 minutes, or somewhat longer than Earth's rate of rotation. The tilt of the axis is 25.2°, or 2° more than Earth's. Like Earth, Mars has seasons, but they are longer, since Mars takes 1.88 years to orbit the Sun. A season is 5.6 Earth months long.

Jupiter's rotation is completed in 9.81 hours and its axis of rotation is inclined by only 3.1°. It is the fastest rotating planet in the solar system, a fact revealed in its shape. If Jupiter is observed through a telescope, it is not disk-shaped; it has a definite oval shape. As Jupiter spins, it develops an equatorial bulge. The faster the spin, the larger the planet, and the less rigid it is (Jupiter is a "gas giant"), the greater the bulge will be. Jupiter's equatorial radius is 71,492 kilometers, whereas its equatorial radius is 66,854 kilometers. (The rocky planet Earth, by contrast, has a polar radius of 6,357 kilometers and an equatorial radius 6,378 kilometers, which results in a difference of only 21 kilometers.)

Another effect of Jupiter's fast rotation is the banding of its atmosphere, with light-colored zones and dark-colored belts. Zones are areas of high pressure and high elevation. They are the upwelling section of a convection cell in the planet's atmosphere. Belts are low-pressure and lower-elevation sections of the cell. Rotation also is instrumental in producing the many ovals and eddies visible in Jupiter's atmosphere. The Great Red Spot, which is larger than several Earths placed side by side, is the largest and oldest of these features. Jupiter's rapid rotation is also responsible for its intense magnetic field. This rapid circulation of the planet's liquid metallic hydrogen core generates this huge field by means of a dynamo effect.

The planet Saturn rotates in 10.63 hours, which is 0.8 hours more than Jupiter's rotation. Since Saturn is also smaller, its equatorial bulge is less prominent than Jupiter's, but an oval outline can still be observed when the planet is viewed through a telescope. Saturn also has banding, but it is much less noticeable than Jupiter's. Saturn's lower gravitational field compresses the atmosphere less than Jupiter's does. Saturn's magnetic field is caused by the rapid rotation of its liquid metallic hydrogen layer, which is similar to Jupiter's.

Furthermore, with an inclination of the spin axis of 26.7°, Saturn should have seasons. One phenomenon that may be caused by the seasonal shift was observed in late 1990. The Hubble Space Telescope photographed Saturn and revealed a large cloud of ammonia spreading across the planet's atmosphere. Scientists speculated that Saturn "burped" ammonia from the atmosphere in response to the seasonal change of sunlight. Similar clouds have been observed since, and they occur at the time of the northern hemisphere's winter. Astronomers expect to learn more about this phenomenon from data returned by the Cassini orbiter.

Uranus poses a major puzzle to astronomers, since it rotates once in 17.18 hours but has an inclination of 97.9°. Scientists speculate that a large object must have collided with Uranus and knocked it on its side. At this time, material was thrown from the planet, with some of it accreting (gradually building up) into the satellites orbiting Uranus. The satellites orbit the planet in the equatorial plane, which indicates that they formed after the collision. Uranus has seasons that are each twenty-one years long, since it revolves around the Sun in eighty-four years. For twenty-one years, one pole is pointed toward the Sun, then for the next twenty-one years the equator faces the Sun, then the other pole, and finally the equator again. Although the planet rotates rapidly, it probably does not have a liquid metallic hydrogen layer. As a result, it does not have a strong magnetic field.

The rotation rate of Neptune is 16.05 hours, with an inclination of 29.6°. Neptune's satellites have strange characteristics; the largest, Triton, has a clockwise direction for its orbit, which is not the norm for the solar system. Nereid has a very elongated orbit, which is also unusual for a satellite. Gravitational interactions between the planet and its satellites may have produced this strange situation.

The dwarf planet Pluto is very small, and its orbit is the most elliptical of the planets. It rotates in 6.387 days and has a 94° inclination. Its orbit takes it closer to the Sun than Neptune for twenty years. Neptune was the ninth planet of the solar system from 1979 until 1999, since Pluto was closer to the Sun during this time. Its slow rotation and unusual inclination could be caused by a gravitational interaction with Neptune and its satellites. Pluto has another oddity: its satellite Charon. Charon orbits Pluto in the equatorial plane, and the two keep the same face toward each other. Charon may be a section of Pluto that was pulled away during a cataclysmic collision.

APPLICATIONS

Orientations and rates of planetary rotation are important to the solar system. Planetary rotation provides clues to events that have occurred in the solar system. From a practical point, the rate and orientation of Earth's rotation are very important to life on this planet. The rate of rotation sets the rhythm of life. It provides periods of light and darkness that govern life's biorhythms. Some creatures use the cover of darkness while foraging, while others (such as lizards) make use of the sunlight to warm themselves. It is difficult to imagine human life on a world where the day is ten hours long, or life on a planet where one face always pointed toward the Sun. If Earth rotated more slowly, perhaps with a day of thirty-six hours, the days would be hotter for a longer period of time. The nights would be colder and longer. A greater temperature difference would, in turn, have produced greater wind speeds and weather in general would be more extreme. Life would have had to evolve to handle these harsher conditions and expend more energy on survival.

Orientation of the rotational axis is also important in determining Earth's seasons. If Earth's inclination were 0°, the climate would vary little. The reason there are such fierce storms in spring is the difference in the temperature of the cold ground and the warming atmosphere. As the Sun warms a hemisphere experiencing a spring season, atmospheric instability and subsequent violent weather result. With zero inclination, the same amount of warming per surface area would occur each day. The area near the equator would be very warm and rainy, while the areas 30° north and south of the equator would be more arid than now with even less rainfall. Polar areas would be colder, and birds and other animals would not migrate as far north for reproduction purposes. On the other hand, if the poles were oriented at 90°, the climate would be very different. For three months, the north polar region would receive sunlight for twenty-four hours a day and would be very warm. The south polar region would be in total darkness and would be very cold. As Earth revolves around the Sun, sunlight would strike the entire surface for twelve hours, and then there would be twelve hours of darkness. Winter for the Northern Hemisphere would follow with continual darkness and cold. Then three months of spring would follow with twelve hours of day and twelve hours of night. Such severe changes over a twelve-month period would be difficult to withstand for life as we know it.

CONTEXT

Recognizing that Earth rotates was a major scientific advance. By looking at the sky, with its moving stars, planets, and the Sun, one gets the impression that the Earth is stationary. The Greeks believed in a geocentric theory of the cosmos, where Earth indeed stood still and all other objects in the cosmos moved around it. This Ptolemaic theory, named for the Greek astronomer Ptolemy, held sway for more than twelve hundred years. It eventually became part of the accepted dogma of Catholicism.

In the mid-sixteenth century, Nicolaus Copernicus stated that Earth and the other planets revolved around the Sun in circular orbits. It took 150 years of work by various scientists for this new "heliocentric" theory of the solar system to be accepted by the scientific community. Also in the sixteenth century, Tycho Brahe compiled the most accurate positional data for the planets. These data were used near the end of the sixteenth century by Johannes Kepler to derive his three laws of planetary motion. Galileo made telescopic observations that verified the heliocentric theory. Sir Isaac Newton provided the theoretical basis for the motion of the planets with his law of universal gravitation. He also used it to reveal the cause of the tides. The Moon and the Sun pull on different parts of Earth with different amounts of force, which places Earth in a tug-of-war that pulls the oceans into the bulges known as the tides.

James Bradley, an English astronomer, discovered the nutation of the Earth in the mid-1700's, when he was making observations of stellar positions. He published his results after nineteen years of careful study. The Coriolis effect was explained by French physicist Gaspard-Gustave de Coriolis, in 1835 when he investigated motion on a spinning surface such as Earth.

Although much work on rotation of the other planets was done before 1957, it was various early spacecraft, such as Mariner 10, the Vikings, the Pioneers, and the Voyagers, that returned data to enable astronomers to refine our knowledge of the planets' rotation. With the Magellan spacecraft sent to Venus, Ulysses and Galileo to Jupiter, Cassini to Saturn, and other planetary spacecraft of the future, scientists will continue to learn more about physical properties of the planets.

Stephen J. Shulik

FURTHER READING

Arny, Thomas T. *Explorations: An Introduction to Astronomy*. 3d ed. New York: McGraw-Hill, 2003. A general astronomy text for the nonscientist. Includes an interactive CD-ROM and is updated with a Web site.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. Filled with color diagrams and photographs, this popular work on solar-system astronomy and planetary exploration takes readers through the Mars Pathfinder and Galileo missions.

Accessible to astronomy enthusiasts and general readers alike.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Includes up-to-date spacecraft information including that from National Aeronautics and Space Administration's (NASA) Great Observatories and rovers on Mars. Includes images that span the electromagnetic spectrum.

Faure, Gunter, and Teresa M. Mensing. *Introduction to Planetary Science: The Geological Perspective*. New York: Springer, 2007. Designed for college students majoring in Earth sciences, this textbook applies general principles of geology to bodies throughout the solar system. Excellent for learning comparative planetology.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college text on astronomy, somewhat more advanced than many introductory texts, with a wealth of detail and excellent diagrams. Chapters 6 through 16 describe the solar system. Comes with a CD-ROM.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Particularly strong in its presentations of Earth-Moon science and the Jovian planets. Takes a comparative planetology approach rather than providing individual chapters on each planet in the solar system.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find standard introductory texts pitched at low a level. Covers all topics from solar-system objects to cosmology.

Leverington, David. *Babylon to Voyager and Beyond: A History of Planetary Astronomy*. New York: Cambridge University Press, 2003. Takes a historical approach to planetary science. Heavily illustrated, concluding with a summary of spacecraft discoveries. Suitable for general readers and the astronomy community alike.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar-system astronomy suitable for an introductory college course but accessible to nonspecialists as well. Filled with supplemental learning aids and solved student exercises. A Web site is available for educator support.

Snow, Theodore P. *The Dynamic Universe*. Rev. ed. St. Paul, Minn.: West, 1991. A general introductory college text on astronomy. Covers historical astronomy, equipment used in astronomy, the solar system, stellar astronomy, galactic astronomy, cosmology, and life in the universe. Featured in the book are special inserts, guest editorials, and a list of additional readings at the end of each chapter.

See also: Earth's Rotation; Planetary Orbits; Planetary Orbits: Couplings and Resonances.

Planetary Satellites

Categories: Natural Planetary Satellites; Planets and Planetology

All but two of the planets in the solar system have one or more smaller bodies in orbit around them. These satellites, popularly called "moons," are important for what they tell scientists about the origins of the planets and the evolution of the solar system.

OVERVIEW

Of the planets in the solar system, all but two, Mercury and Venus, have at least one satellite. There are more than one hundred known satellites in the solar system, ranging in size from about 15 kilometers to more than 5,200 kilometers in diameter; the number continues to grow as spacecraft and telescopic observations discover more and more relatively tiny satellites about the gas giants, particularly Jupiter and Saturn. Although the largest satellites of the solar system's eight planets have been identified, a table of known satellites would become

Data on Major Satellites in the Solar System

Planet	*Satellite*	*Diameter (km) or Dimensions*	*Mass (kg)*	*Density (g/cm^3)*
Earth	Moon	3,476	7.35×10^{22}	3.34
Mars	Phobos	27 × 22 × 19	9.60×10^{15}	2
	Deimos	15 × 12 × 11	1.90×10^{15}	2
Jupiter	Io	3,630	8.92×10^{22}	3.53
	Europa	3,138	4.87×10^{22}	3.03
	Ganymede	5,262	1.49×10^{23}	1.93
	Callisto	4,800	1.08×10^{23}	1.70
Saturn	Mimas	392	4.50×10^{19}	1.43
	Enceladus	500	7.40×10^{19}	1.13
	Tethys	1,060	7.40×10^{20}	1.19
	Dione	1,120	1.05×10^{21}	1.43
	Rhea	1,530	2.50×10^{21}	1.33
	Titan	5,150	1.35×10^{21}	1.89
	Iapetus	1,460	1.88×10^{21}	1.15
	Phoebe	220	?	?
Uranus	Miranda	480	7.50×10^{19}	1.26
	Ariel	1,330	1.40×10^{21}	1.14
	Umbriel	1,110	1.30×10^{21}	1.82
	Titania	1,600	3.50×10^{21}	1.60
	Oberon	1,630	2.90×10^{21}	1.28
Neptune	Triton	3,800	1.30×10^{23}	?
	Nereid	300	2.10×10^{19}	?
	1989 N1	400	?	?
Pluto	Charon	1,150?	?	2 ?

Note: Jupiter, Saturn, Uranus, and Neptune also have other, minor satellites. Venus and Mercury have no known satellites. Pluto, designated a dwarf planet by the International Astronomical Union in 2006, is considered by some to form a double-planet system with Charon.

out of date within a matter of months; more are being discovered all the time. Satellites are important scientifically for the clues they provide to the origin of their parent planets and the solar system in general.

Most models of planetary formation start with some variation on the "nebular hypothesis" devised by German philosopher Immanuel Kant in 1755 and later modified in 1796 by French mathematician Pierre-Simon Laplace. This hypothesis in modern form suggests that the Sun, planets, satellites, and smaller debris in the solar system started as a rotating flattened cloud of gas and dust. This rotating cloud eventually became unstable and broke into rings. Material concentrated in the bulging center of the cloud became compressed, eventually giving birth to the Sun. The gaseous rings cooled eventually, allowing precipitation of solid crystals that aggregated to form rocky materials. These aggregates, in turn, accreted over geological time into larger and larger bodies. The largest of these bodies became the major planets, with smaller bodies generally falling into the planets, adding to their mass. Under certain circumstances, small bodies were captured as satellites or remained freely orbiting the Sun as dwarf planets, asteroids, and comets.

Like the major planets, satellites vary in composition depending on their mean distance from the Sun. Satellite composition can be expressed in terms of two major components: ice and rocky silicate material. Satellites in the inner solar system (Earth's moon, for example) are composed almost exclusively of rocky material, whereas those orbiting Jupiter and beyond are mixtures of ice and rocky material. Thus, satellite density decreases in the outer solar system. The distribution of ice and rock among satellites parallels the preponderance of rocky planets in the inner solar system (Mercury through Mars) and gas-rich planets in the outer solar system (Jupiter through Neptune). Pluto is composed almost completely of ice and has been reclassified as either a dwarf planet or plutoid, the first member of a class of objects from the Kuiper Belt referred to as plutinos. From this distribution, it can be noted that "volatile" substances (those with low melting and boiling points) are concentrated in the outer solar system. "Refractory" substances, those with high melting and boiling points, occur in greater abundance within the inner solar system. Satellites in the solar system reveal a considerable diversity in terms of surface composition and physical characteristics.

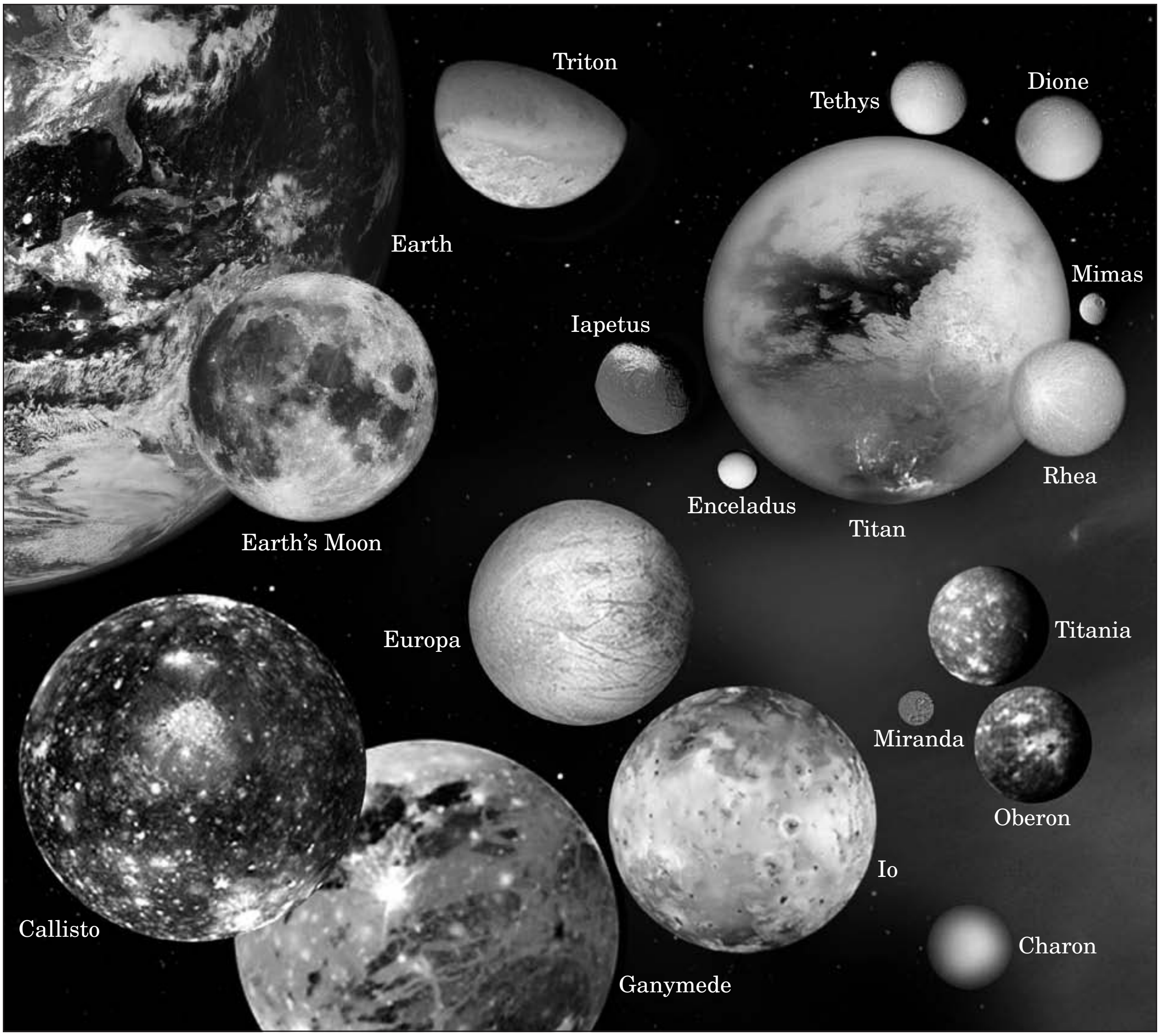
Earth's moon, the nearest satellite to the Sun, is a refractory-rich rocky triaxial ellipsoid approximately 3,500 kilometers in average diameter, more than one-fourth the size of Earth. The Moon is the best-known of all other extraterrestrial satellites because of the many spacecraft that have examined it from orbit and landed on its surface. These robotic missions began with the 1959 mission by the Soviet Union to photograph the lunar "far side," followed by the National Aeronautics and Space Administration's (NASA's) successful Ranger, Surveyor, and Lunar Orbiter projects in the 1960's. The most famous lunar project of course was NASA's Apollo program, which landed humans on the Moon. From 1969 to 1972, six Apollo missions landed twelve astronauts on the Moon, who returned numerous rocks and "soil" samples. The Apollo astronauts also set up experimental equipment for measuring "moonquakes" and other phenomena. Information from all the lunar missions, whether crewed or robotic, indicate that the Moon is a complex, small satellite with a metallic core, an iron-rich silicate mantle, and a silicon plus aluminum-rich silicate crust.

Nearly two decades transpired before another spacecraft was sent to the Moon after the Russian Luna 24 mission returned a small amount of rock and soil samples to Earth from the Moon's Sea of Crises in 1976. By this point, NASA had moved on to attempting to land robotic spacecraft, the Vikings, on Mars. In early 1998 NASA sent the Lunar Prospector into lunar orbit. Outfitted with five different instruments, Lunar Prospector's objectives were to globally map the lunar resources, determine the Moon's complex gravity field and its minor magnetic field, and to look for any outgassing from the surface. The surprising result of Lunar Prospector investigations was the strong indication that between 10 to 100 million metric tons of water ice could be located in permanently shadowed portions of craters near the lunar poles.

The presence of water on the Moon as a resource that could be utilized would greatly enhance the viability of crewed lunar research outposts, especially if those outposts were constructed at or very near the Moon's poles. This finding was not direct evidence. The spacecraft's neutron spectrometer picked up the presence of protons; the most reasonable extrapolation of the data was that those protons were bonded together in water ice. The potential finding of water on the Moon was not a total surprise. Two years earlier, a Ballistic Missile Defense Organization probe called Clementine used the Moon as a testbed for new sensor technologies, among other things. Some of the sensor data hinted at the possibility of water near the poles. Lunar Prospector took the search for water to a new level, and the real surprise was the total amount of water that Lunar Prospector seemed to find.

The Moon has no significant atmosphere, a trait it shares with most other small bodies in the solar system (Saturn's satellite Titan and Neptune's satellite Triton are exceptions). Also, the Moon's crust is far less complex than Earth's. It consists of densely cratered highland areas composed of a feldspar-rich (calcium-aluminum silicate) rock called anorthosite (light-colored areas), with dark basalt lava flows (iron-rich silicate rock) filling huge impact craters that the Italian astronomer Galileo named *maria* (Latin for "seas"). The Moon contains no real granite rocks such as those that compose much of Earth's continents.

The Moon always shows the same face to Earth as it orbits, because the Moon's rotational rate is equal to its orbital rate. This situation is an example of "synchronous rotation," in which the rotation rate of a body has some precise mathematical relationship to orbital period (time required to complete one orbit). The Moon's 1:1 ratio of orbital-to-rotational period results from the Earth-side of the Moon bulging out because of gravitational tidal forces between Earth and the Moon. Eventually, this bulge (the side facing Earth) comes to lie along the Earth-Moon line, its most stable configuration. Other bodies in the solar system show similar relationships. For example, the satellites of Mars, Phobos and Deimos, rotate so that the same side always faces Mars. Many satellites of Jupiter and Saturn also show this relationship.



Moons of the solar system's planets appear here in composite. (NASA)

Pluto and its satellite Charon both revolve and rotate at the same rate. This means that an observer on Pluto would always see Charon in precisely the same place in the sky all day and night.

Traveling out from the Sun, the next planet is Mars, with its tiny satellites Deimos and Phobos. These oddly shaped, rocky bodies (neither is spherical) are most likely escapees from the nearby asteroid belt, a zone between Mars and Jupiter that contains thousands of small planetoids (up to about 1,000 kilometers in diameter), rock fragments, and dust. Photographed up close by the Viking Orbiter 1 in 1977, Phobos and Deimos appear to be composed of the same materials that occur in certain meteorites.

At Jupiter, a miniature solar system is found. Jupiter and most of satellites were photographed extensively by Voyagers 1 and 2 in 1979 and earlier by Pioneer 10. Jupiter's four largest satellites are called the "Galilean satellites" because they were discovered by Galileo in 1610. Their densities and rock-to-ice ratios decrease with increased distance from Jupiter, a relationship that is mirrored by the larger solar system. In addition, geological activity decreases from the closest Galilean moon, volcanically active Io, to the highly cratered outermost satellite, Callisto. The high crater density on Callisto's surface shows that its surface is very old and has not become "resurfaced" by high-energy processes such as volcanic activity or erosion. This same reasoning is used elsewhere in the solar system to deduce relative ages of satellite and planetary surfaces. Earth's moon, Mercury, and many other bodies are heavily cratered by impacting projectiles and, therefore, are considered to have old surfaces. The Galilean satellites Io and Europa show no craters because their surfaces are renewed constantly by molten sulfurous compounds on Io and liquid water that freezes to ice on the surface of Europa. The heat source to produce this volcanic activity results from tidal forces originating from massive Jupiter.

Saturn and most of its sixty satellites were photographed and studied by Pioneer 11, the two Voyager spacecraft, and the Cassini orbiter. Most of its satellites are icy, heavily cratered worlds; one satellite, Titan, is the second largest satellite in the solar system (Jupiter's Ganymede is larger). Titan appears to have a surface composed of complex hydrocarbon compounds and liquid nitrogen, but it is unusual for such a small "planet" to have its extensive atmosphere, which is composed mostly of methane (CH_4). Its presence may be caused by the extremely cold temperatures or by the continual production of gases by some type of its cryogenic volcanic activity.

Uranus, visited only by Voyager 2 in 1986, has at least twenty-four satellites, of which only five were large enough to have been observed from Earth prior to Voyager's visit; many were discovered long after that visit. Of all the satellites photographed by Voyager 2, the most surprising by far is Miranda. Miranda has nearly crater-free dark areas with concentric grooves and chevron-shaped features that are separated by huge fault scarps (cliffs) from heavily cratered areas. Proposals to explain Miranda's bizarre surface features included the breakup and reassembly of the satellite following a catastrophic collision with another body or the upwelling of heated water to form its dark, bulging, grooved terrains.

Neptune was visited in 1989 by Voyager 2 and its two major satellites, Nereid and Triton, were photographed, along with six new ones. One of those, 1989 N1, measures about 400 kilometers in diameter, replacing Nereid as the second largest Neptunian satellite. Nevertheless, Triton is the most important. It shows a frozen surface with virtually no impact craters. Triton's landscape includes huge frozen "lakes" of solidified water mixed with ammonia, along with a so-called cantalouped terrain of intersecting grooves and ridges that may represent fault systems. Frigid Triton has a surface temperature of only 37 kelvins. It has scattered, dark linear streaks on its surface that may represent nitrogen-powered geysers spewing dark organic matter out on the surface. Winds blowing in Triton's thin methane atmosphere align these geyser streaks and blow thin, ice crystal clouds around this miniature planet. Triton is a unique world that has somehow maintained internal heating and recent resurfacing. It is truly a remarkable feat for a body in such a cold place.

APPLICATIONS

Satellites are studied primarily for what they tell scientists about the origin of the solar system. As intelligent beings who are products of solar system evolution, humans have a natural curiosity about the solar system's and our own origin. The mechanisms of satellite formation and eventual capture by larger planets illustrate many of the dynamic processes that have shaped the solar system since its condensation from the solar nebula perhaps 4.5 billion years ago.

As an example, Earth's moon poses interesting problems regarding attempts to understand the origin of the Earth. Many ideas have been proposed for the origin of the Earth-Moon system. These ideas include the hypothesis that the Moon formed elsewhere in the solar system and was later captured by the Earth, or that the Moon "fissioned" off from the early Earth itself, possibly leaving a major hole how known as the Pacific Ocean. Detailed analyses of lunar samples returned by Apollo astronauts (1969-1972) and the Soviet robotic Luna missions (which concluded in 1976) show that lunar rocks are similar in some important respects to terrestrial rocks, but they also show critical differences. Lunar and earthly rocks have the same ratios of the three oxygen isotopes (nuclei with the same atomic numbers but different mass number—hence different amounts of neutrons but the same number of protons): oxygen 15, oxygen 17, and oxygen 18. This indicates that their constituents condensed from the same area in space. Thus, the theory that postulates the Moon was captured after forming elsewhere is no longer held in much esteem.

The fission theory contrasts with the idea that the Moon was formed at some area removed from the Earth. Concentrations of other chemicals in lunar samples than oxygen isotopes display great differences from Earth rocks. For example, the Moon is richer in iron and carbon compared to Earth and contains far more refractory elements (such as titanium and calcium) compared to volatile substances (such as water, sodium, and potassium). These differences indicate that Earth and the Moon could not have formed from exactly the same starting materials as would be expected if they condensed from a common area of the original solar nebula. Therefore, the fission theory is also no longer considered viable.

An idea to resolve the problem of lunar origin, called the "impactor hypothesis," involves a violent collision of the early Earth with a wayward, Mars-sized planet (the "impactor"). This collision would have blasted out material from Earth all the way to the metallic core and would volatilize or pulverize most of the impactor. A mixture of loose impactor and Earth materials then would have orbited the wounded Earth, eventually accreting (accumulating) to form the Moon. Advocates of this hypothesis argue that mixing of impactor and Earth materials to form both the Moon and Earth would explain the similar oxygen isotope ratios. Yet, because the Moon and Earth would be composed of different ratios of proto-Earth and impactor materials, the two bodies' differerces in major chemical components such as water and iron are also explained. Fortunately for the future evolution of life, Earth acquired most of the volatiles, including water, which is essential to initiate and sustain life. This theory presently is preferred.

The complexity of planet and satellite formation is well illustrated by bodies in the outermost reaches of the solar system; the Neptunian system and the Pluto-Charon system are particularly intriguing. Large, icy Triton orbits Neptune in a "retrograde" direction, clockwise when viewed from "above" the ecliptic plane of the solar system. Most planets and satellites rotate and revolve in a counterclockwise direction. Voyager 2 images and measurements show that Triton and Nereid are very similar to Pluto-Charon in composition and density. All are methane-rich ice balls that resemble the comets that originated in the outer solar system. Theorists speculate that Triton, Nereid, and perhaps Pluto may be comet-like objects that were captured by Neptune early in the history of the solar system and revolved around Neptune in the normal direction. Triton's retrograde orbit may have resulted from a collision or close encounter with a large passing planetoid, an event that may also have caused Nereid to assume its strange, elongated elliptical orbit. In this model, Pluto was split in two (to make Charon) by the encounter and ejected into the outer solar system. Regardless of whether this hypothesis is correct, the Neptunian system is known to be unstable, indicating that something has disturbed it since the original formation of the solar system. In about 10 to 100 million years, Triton will spiral close enough to Neptune to be torn to pieces by tidal forces, adding greatly to the mass of Neptune's current thin system of partial arcs.

Although most study of planetary satellites is conducted to learn more about the solar system in general and the possible future of Earth in particular, knowledge of one nearby satellite, Earth's moon, will eventually be used to practical advantage. For example, dark lava flows (basalt) on the Moon contain abundant titanium, a valuable component in high-temperature metal alloys (like metal used in rocket bodies). Rocks in the lunar highland areas are mostly anorthosites, composed primarily of the mineral plagioclase feldspar, which is a potential source of aluminum. Some aluminum is extracted from feldspar deposits on Earth, but the process to separate aluminum from the rest of the mineral requires a high-energy input. On Earth, this process is accomplished commonly by locating the aluminum smelter near a source of hydroelectric power; on the Moon, the energy would have to come from the Sun or a small local nuclear power plant.

Earth's Moon is a potential base for launching spacecraft to other regions of the solar system. This endeavor would require the construction of a lunar base where the Moon's low gravity would greatly facilitate spacecraft launches. Far less energy must be expended on the Moon to escape its gravity field than is required on Earth. Wholesale colonization of the Moon to alleviate population pressures on Earth, however, is not feasible. One aspect of the scientific study of the Moon confirms the distinct lack of concentrated water supplies on the Moon, although some researchers believe that subsurface water in localized "permafrost" deposits will be discovered in the future, such as those found on Mars. Even if they exist, these frozen water deposits would not support large populations of humans.

In 2004, in the aftermath of the *Columbia* accident, President George W. Bush charged NASA with fulfilling his administration's Vision for Space Exploration, which called for a permanent human outpost on the Moon and development of technology to expand the human presence to Mars and beyond. NASA pursued the Constellation program as a versatile crewed transportation system to return astronauts to the Moon, build a base there near the South Pole where the water appears to be located in greatest abundance, and then work toward eventual trips to Mars. A nonbinding deadline for reaching the Moon was set at 2020.

However, the United States was not the only nation with an interest in sending humans back to the Moon. The Chinese had already announced their plans for sending taikonauts (the Chinese word for astronaut) into space and in due course to the Moon. The Europeans sent the Small Missions for Advanced Research in Technology 1 (SMART 1) spacecraft to the Moon to produce the highest-resolution mapping of its surface. A Chinese satellite named Chang'e 1 was sent into orbit, the Japanese sent their Kaguya probe to the Moon, and the Indians launched their Chandrayaan spacecraft on October 22, 2008. For the first time an international fleet of probes was investigating the Moon, and with a government-supported space program, it appeared that Chinese taikonauts might reach the Moon by the second decade of the twenty-first century.

CONTEXT

The five planets known to ancient peoples were referred to as the "wandering stars" to distinguish them from the "fixed stars." At least by the time of the ancient Greeks, the Moon was recognized as a satellite of the Earth. Most Greek philosophers, such as Plato and Aristotle, believed that all planets and the Sun revolved around the Earth.

The history of astronomy and of science in general was influenced profoundly by Galileo's discovery of the Jovian satellites. Galileo had constructed a crude telescope and used it to scan the heavens. He discovered the four largest satellites of Jupiter (Io, Europa, Ganymede, and Callisto) now known as the "Galilean satellites" in his honor. In 1610, Galileo published his findings in *Sidereus Nuncius* (starry messenger).

Before that, the heliocentric (sun-centered) model of the solar system proposed by Nicolaus Copernicus was seriously questioned and indeed considered heretical by the Catholic Church. Galileo's discovery of a "miniature solar system" consisting of satellites around a planet demonstrated that objects could revolve around something other than the Earth, thus displacing the Earth from the center of the universe.

The next major discovery—the large satellite of Saturn, Titan—was made by the Dutch astronomer Christiaan Huygens in 1656. By the end of the nineteenth century, eight more satellites had been discovered orbiting Saturn, many of them found by the Italian astronomer Gian Domenico Cassini. Huygens also realized that the "ears" on either side of Saturn first described by Galileo were actually rings, a feature attributed to individual particles (tiny satellites) by James Clerk Maxwell in 1857.

Although the distant planet Uranus was discovered in 1781 by English astronomer Sir William Herschel, its five largest satellites required an additional 167 years to detect. The last one, tiny Miranda, was discovered by American astronomer Gerard Peter Kuiper in 1948.

Neptune was discovered in 1846 by German astronomer Johann G. Galle, followed in the same year by the discovery of its large satellite Triton by English astronomer William Lassell. Charon, the largest satellite of Pluto, was discovered in 1978 by James W. Christy of the United States Observatory after he noticed that a photographic image he had taken of Pluto showed a lump on one side. This lump was shown later to move relative to Pluto, confirming the existence of a satellite.

These later discoveries of other satellite systems, along with continuing new discoveries, have proven crucial to the understanding of the origin of the planets, as well as the satellites themselves. Scientific study of satellites shows that the solar system evolved amid violent collisions, gravitational perturbations of orbits, and heating and cooling of surfaces and interiors, in a bewildering variety of combinations. These studies show that, although they share some characteristics, every planet-satellite system formed in some unique way according to conditions prevailing in its particular region of the solar system.

John L. Berkley

FURTHER READING

Fischer, Daniel. *Mission Jupiter: The Spectacular Journey of the Galileo Spacecraft*. New York: Copernicus Books, 2001. Suitable for a wide range of audiences. Thoroughly explains all aspects of the science and engineering of the Galileo spacecraft. Particularly good are the discussions about the nature of the Galilean satellites.

Greenberg, Richard. *Europa the Ocean Moon: Search for an Alien Biosphere*. New York: Springer, 2005. A complete description of current knowledge of Jupiter's satellite Europa through the post-Galileo spacecraft era. Discusses the astrobiological implications of an ocean underneath Europa's icy crust in a readable text geared for astronomy enthusiasts as well as college students. Well illustrated.

Harland, David M. *Cassini at Saturn: Huygens Results*. New York: Springer, 2007. The text provides a thorough explanation of the entire Cassini program, including the Huygens landing on Saturn's largest satellite, Titan. Essentially a complete collection of NASA releases from the start of Cassini flight operations through the majority of Cassini's seventy orbits of its primary mission, which concluded a year after this book entered print. Somewhat technical.

_______. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002. A technical description of the Cassini program, its science goals, and the instruments used to accomplishment those goals. Written before Cassini arrived at Saturn, it provides a historical review of pre-Cassini knowledge of Saturn and its satellites.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college-level textbook written in a style accessible to nonscience majors. New terms are written in boldface type and defined in the text. Lavishly illustrated with well-executed drawings and black-and-white photographs. Hartmann is a distinguished planetary scientist and an accomplished artist. His paintings of remote planetary scenes have fostered a greater understanding of the variety of features displayed by planetary surfaces. Includes appendices, with extensive planetary vital statistics, a chapter-by-chapter bibliography, and detailed index.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs.

Lovett, Laura, Joan Harvath, and Jeff Cuzzi. *Saturn: A New View*. New York: Harry N. Abrams, 2006. A coffee-table book replete with about 150 of the best images returned by the Cassini mission to Saturn. Covers the planet, its many satellites, and the complex ring systems.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar system astronomy suitable for an introductory college course, but accessible to nonspecialists as well. Filled with supplemental learning aids and solved student exercises. A Web site is available for educator support.

Rothery, David A. *Satellites of the Outer Planets: Worlds in Their Own Right*. New York: Oxford University Press, 1999. Up to date with the Galileo spacecraft results, this heavily illustrated, easy-to-read book offers a comprehensive geological study of the satellites of the gas giants.

Van Pelt, Michel. *Space Invaders: How Robotic Spacecraft Explore the Solar System*. New York: Springer, 2006. A historical account of robotic planetary science missions attempted by all spacefaring nations, written by an European Space Agency cost and systems engineer. The narrative not only explains the science but also provides a behind-the-scenes description of the development of a space exploration mission from concept proposal to flight operation.

See also: Callisto; Enceladus; Europa; Ganymede; Iapetus; Io; Jupiter's Satellites; Jupiter's Ring System; Lunar History; Mars's Satellites; Miranda; Neptune's Satellites; Neptune's Ring System; Planetary Formation; Pluto and Charon; Saturn's Satellites; Titan; Triton; Uranus's Rings; Uranus's Satellites.

Planetary Tectonics

Category: Planets and Planetology

A tectonic process is any process that causes movement or distortion of a solid planetary surface. The gas giant planets do not have solid surfaces, so tectonics does not apply to them. The other planets and satellites in the solar system, however, have varying types of tectonic activity, which can include volcanic activity, activity or movement along fault lines, uplifting, and crust folding.

OVERVIEW

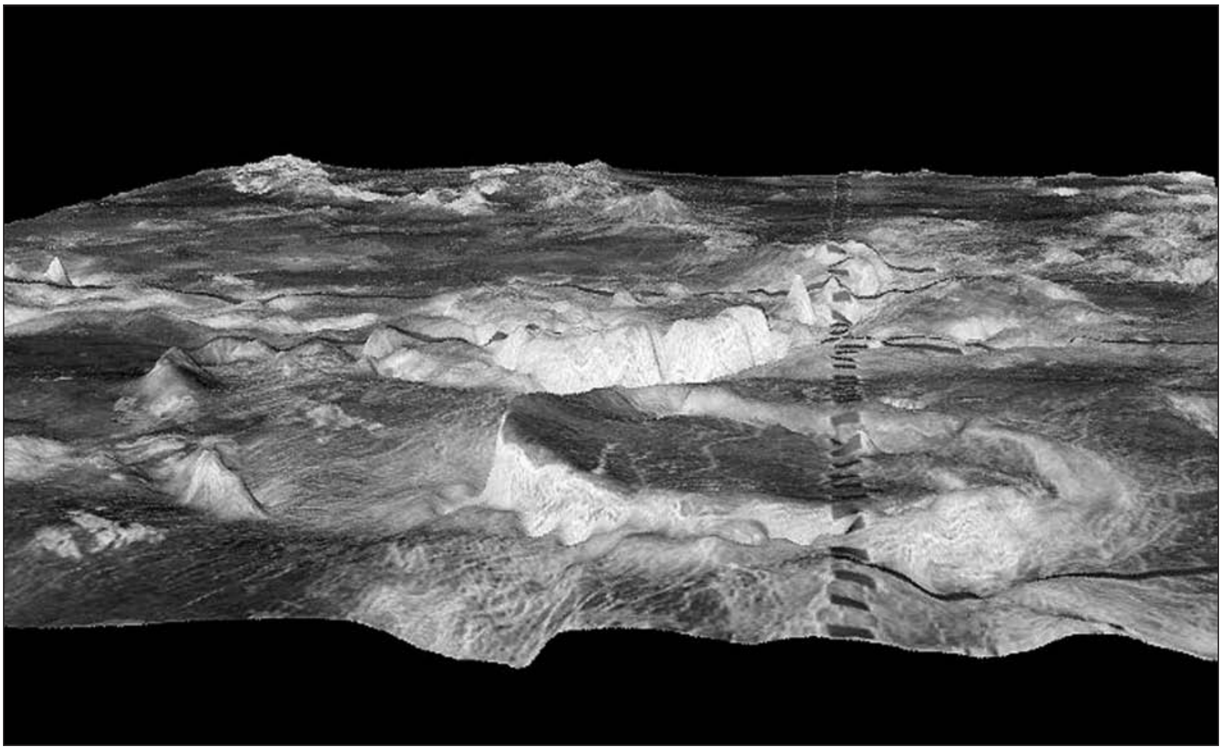
The contours of the eastern coasts of North and South America fit nicely into those of the western coasts of Europe and Africa, suggesting that these continents were at one time a single landmass. Various geological and fossil patterns also match. On the basis of this evidence, Alfred Wegener proposed the theory of "continental drift" in 1912. At the time, few geologists took Wegener's ideas seriously; the idea that continents might move, even if only a few centimeters a year, seemed too far-fetched. Furthermore, Wegener was unable to suggest a mechanism that might cause this motion.

Evidence favoring Wegener's ideas continued to accumulate, however. By the 1960's, geologists had begun to embrace the theory of plate tectonics. On Earth the crust is divided into several large plates that float on the mantle, which is in a plastic-like state between liquid and solid. Slow convection currents in the mantle cause the plates to drift very slowly. Hence, the theory of plate tectonics confirms Wegener's idea of continental drift, provides a mechanism for the drift, and explains Earth's geological structures and evolution. The theory of plate tectonics was the major geological advance of the twentieth century.

On Earth, most earthquakes and volcanic activity occur at the boundaries between plates. Plate boundaries are of three types: divergent, where plates move apart; convergent, where plates move together; and transform, where plates slide horizontally past each other.

Plates diverge along the ocean ridge-rift system. An example is the Mid-Atlantic Ridge, which runs roughly north-south along the floor of the Atlantic Ocean midway between Europe and Africa, to the east, and North and South America, to the west. As the plates on each side of the Mid-Atlantic Ridge diverge, hot magma wells up from below and oozes out along a series of fissures, adding new ocean crust to each plate. The new crust is elevated into an undersea ridge with a central rift. At spots where the volume of extruded lava is especially large, volcanic islands such as Iceland can build up above sea level.

There are three types of convergent boundaries: oceanic-continental, oceanic-oceanic, and continental-continental. Where an ocean plate collides with a continental plate, such as along the west coast of South America, a subduction zone forms. The oceanic plate, composed of slightly denser rock, subducts (descends) below the continental plate, sinking back into the mantle. The continental plate remains on the surface, sliding over the subducting ocean plate, and its leading edge is crumpled and buckled into a mountain range, such as the Andes. When two oceanic plates converge, one of them subducts under the other, and a volcanic island arc develops along the upper oceanic plate. Subduction zones (at both oceanic-continental and oceanic-oceanic boundaries) are marked by deep-sea trenches, formed where the descending oceanic plate bends downward. The deepest of these is the Mariana Trench in the western Pacific Ocean, which drops to more than 11 kilometers below sea level. When two continental plates collide, neither plate subducts. Instead,

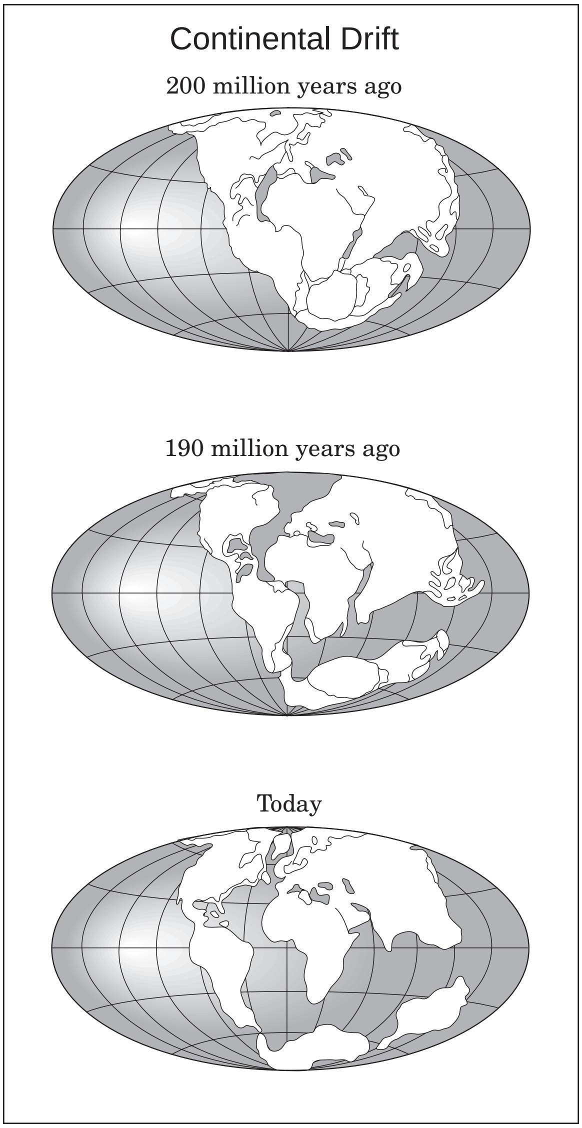


Venus's Galindo quadrangle, with Atete Corona in the foreground. Coronae are a good example of volcanic features formed on Venus from the crust's vertical tectonic motion. (NASA/JPL/USGS) the leading edges of both plates are compressed and crumpled, forming along the suture a high mountain range, such as the Himalayas, which developed when India collided with the rest of Asia.

Most transform boundaries are short and offset segments of divergent boundaries. A few transform boundaries are long, like the San Andreas fault, which cuts across California from its border with Mexico to north of San Francisco. The land to the west of the fault is part of the Pacific Plate, and it is sliding northwest past the North American Plate, which forms the eastern side of the fault.

Earth is the only planet with plate tectonics, but other planets and satellites have had different types of tectonic activities. As a general principle, planets or satellites with very heavily impact-cratered surfaces have had relatively little tectonic activity. The heaviest era of crater formation in the solar system was shortly after its formation, when there was still plenty of debris to crash into solid planets or satellites to form impact craters. If a planet or satellite renews its surface by some type of tectonic activity, the renewal obliterates the craters. Hence, a planetary surface with many impact craters has not been renewed and has undergone little tectonic activity, whereas one that shows evidence of surface renewal may have been subject to tectonic activity.

Both the Moon and Mercury are covered with impact craters, telling us that they have had relatively little tectonic activity in their geologic histories. This fact is related to their relatively small sizes. Smaller objects cool faster; a spoonful of hot soup, for example, cools faster than a large bowl. Small worlds have less interior heat to drive tectonic activity. Both Mercury and the Moon have fractures and fault valleys that formed from meteorite impacts. On the Moon, basins that formed from large impacts were partially filled in by lava flows to form the maria, or so-called “seas.” Rilles, such as Hadley’s Rille near the Apollo 15 lunar landing site, were once rivers of flowing lava. The lunar lava flowed much more easily than lava on Earth, so it did not form extensive volcanoes. Mercury, by contrast, has intercrater plains, rather than maria; these plains probably are lava flows from billions of years ago. Mercury’s geologic history is, however, less well understood than the Moon’s. Mercury also has many scarps, which are cliffs stretching for hundreds of kilometers. They probably formed when Mercury’s interior cooled and shrank, causing the crust to buckle.



Venus has an extremely large number of volcanoes. There are approximately sixteen hundred large volcanoes on Venus and perhaps as many as hundreds of thousands of smaller volcanic features. Most of the large volcanoes are shield volcanoes, which are similar to those on the Hawaiian Islands. Shield volcanoes form when lava flows up through a tube in the center of the volcano. Venus has much larger volcanic features called coronae (the singular is corona) and smaller volcanic structures called either pancake or volcanic domes. Both of these types of structures formed when hot lava flowed up from the mantles and pushed the crust upward, causing it to swell. In addition to these volcanic features, about 80 percent of Venus's surface is covered with solidified lava flows. Geologic evidence suggests that the surface of Venus was renewed by tectonic activity as recently as a few hundred million years ago. Venus does not have the types of tectonic features that are associated with boundaries between tectonic plates on Earth and thus Venus does not have plate tectonics, as Earth does. Rather than causing crustal plates to move horizontally as on Earth, convection currents in Venus's mantle cause the crust to move vertically.

Mars has what are probably the most impressive tectonic features in the solar system. The Valles Marineris is a large tectonic canyon on Mars formed by a crack in the crust that is roughly as deep as the Grand Canyon. It is, however, a few hundred kilometers wide, about as wide as the Grand Canyon is long. It is long enough that, if the Valles Marineris were placed in the continental United States, it would stretch from about Cape Hatteras, North Carolina, to Los Angeles. Valles Marineris probably formed when tectonic uplifting from the Martian interior cracked the crust. Mars's volcanic mountains are equally impressive. Olympus Mons, which is the largest volcanic mountain known in the solar system, towers about 25 kilometers above the Martian surface. Its base is about 700 kilometers wide. The Tharsis Ridge volcanoes are almost as large. These Martian volcanoes are shield volcanoes but are much larger than shield volcanoes on Earth, precisely because Mars does not have plate tectonics. Shield volcanoes form above volcanic hotspots. On Earth the plates drift, while the hotspots remain stationary. During dormant periods on Earth, the portion of the drifting tectonic plate above the hotspot moves. A chain of smaller volcanic mountains results. On Mars, however, the plates do not drift during dormant periods, so the volcano repeatedly erupts at the same location and a single, huge mountain results. Martian volcanoes are now extinct.

Lacking solid surfaces, the Jovian planets do not have tectonic activity, but some of the Jovian satellites do. For example, Io, the innermost satellite of Jupiter, is perhaps the most volcanically active world in the solar system, with about sixty active volcanoes observed by the Galileo orbiter. Io also has lava flows and other features associated with volcanic activity. The volcanoes on Io are distributed randomly rather than along tectonic plate boundaries, suggesting that Io does not have plate tectonics similar to Earth's. Jupiter's second moon, Europa, has an icy crust covering a water mantle. The crust has many cracks caused by shrinking as the watery layers melted. The tectonic activity on both Io and Europa is largely driven by tidal forces from Jupiter heating their interiors. There is also volcanic activity on Saturn's Enceladus and Neptune's Triton.

KNOWLEDGE GAINED

Wegener's continental drift theory was held in poor regard by geologists for several decades. By the 1960's, however, undersea mapping had revealed the Mid-Atlantic Ridge, a chain of undersea volcanic mountains marking the plate boundaries. Scientists also discovered matching patterns of magnetism on either side of the ridge. These lines of evidence, which were not available in Wegener's time, led to the theory of plate tectonics, which encompassed and provided a mechanism for the continental drift theory.

With the advent of the space age, exploration of the solar system provided evidence of tectonic activity on other worlds. The Apollo missions sent astronauts to directly study the lunar surface and bring back samples for additional study. Hence our knowledge of the Moon is second only to our knowledge of Earth. Radar maps of Venus made from Earth using the Arecibo radio telescope and from spacecraft orbiting Venus revealed the considerable volcanic activity on Venus. The most detailed radar maps of Venus were made by the Magellan spacecraft in the early 1990's. The impressive tectonic features on Mars were first discovered by the early Mariner missions, launched in the late 1960's.

Each subsequent successful mission to Mars has increased our understanding of the tectonic forces that shaped the Martian surface.

The Voyager missions to the outer solar system discovered many tectonic surprises including the tectonic and volcanic activity on Europa and Io. In addition, Neptune's largest satellite, Triton, is so cold that it has volcanic and tectonic activity related to solid and liquid nitrogen. Smaller satellites should have cooled too much to be tectonically active, but Saturn's 500-kilometer-diameter satellite Enceladus shows signs of tectonic activity discovered by the Voyager mission and further studied by the later Cassini mission. As humanity continues to explore the outer solar system, it is likely that more unexpected evidence of tectonic activity will be discovered.

CONTEXT

Because we live on Earth and have easy access to its tectonic features, planetary scientists understand tectonic activity on Earth more fully than for any other world in the solar system. Even the best photos from a spacecraft cannot substitute for a geologist's ability to stand on and explore one of Earth's many tectonic features. This knowledge of tectonics on Earth, however, helps planetary scientists understand and interpret tectonic features that spacecraft have revealed on other planets. Similarly planetary scientists who discover new types of tectonic features on other worlds and study how tectonic processes work on other worlds help geologists understand tectonic activity on Earth.

Tectonic activity on Earth often manifests itself in natural disasters such as earthquakes and volcanoes. With increased understanding of planetary tectonics, it may eventually be possible to predict when tectonic events will occur on Earth, or at least to forecast the conditions that may lead to these disasters. The ability to predict earthquakes, volcanic eruptions, and tsunamis caused by undersea tectonic events could save many lives, just as advances in atmospheric science and computer modeling have increased our ability to predict and track major storms and hurricanes, which now saves many lives. Space exploration has contributed to this ability to predict major storms and may similarly make it possible to predict natural disasters caused by tectonic events.

Paul A. Heckert

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Chapters 6 through 10 of this very readable introductory astronomy textbook cover the terrestrial planets, including good discussions of the tectonic activity shaping the surfaces of planets and satellites.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. Chapter 11 of this introductory astronomy textbook is a complete, readable overview of Mercury, Venus, and Mars, including their tectonic activity.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. Chapter 10 of this textbook on the moons and planets of the solar system summarizes our understanding of volcanic and other tectonic processes on both the terrestrial planets and various moons.

Hester, Jeff, et al. *Twenty-first Century Astronomy*. New York: W. W. Norton, 2007. Chapters 6 and 7 of this well-illustrated astronomy textbook are about the terrestrial planets. Tectonic processes are well explained, and the comparison of these processes on different planets is very good.

Morrison, David, Sidney Wolf, and Andrew Fraknoi. *Abell's Exploration of the Universe*. 7th ed. Philadelphia: Saunders College Publishing, 1995. The terrestrial planets are covered in chapters 12 through 15, and the satellites of the outer planets are covered in chapter 17 of this classic astronomy textbook.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. Cambridge, England: Cambridge University Press, 2002. An extremely well written introductory astronomy textbook. Chapter 9 is an overview of the terrestrial planets; the satellites of the Jovian planets are covered in chapter 10.

Zeilik, Michael, and Stephen A. Gregory. *Introductory Astronomy and Astrophysics*. 4th ed. Fort Worth, Tex.: Saunders College Publishing, 1998. Aimed at undergraduate phys-

ics and astronomy majors, this textbook goes into more mathematical depth than most introductory astronomy textbooks. Chapters 4 and 5 cover the basic principles of the terrestrial planets, including tectonic processes.

See also: Earth System Science; Earth's Age; Earth's Composition; Earth's Core; Earth's Differentiation; Earth's Magnetic Field: Secular Variation; Earth's Mantle; Earth's Rotation; Earth's Structure; Ganymede; Io; Mars's Volcanoes; Planetary Interiors; Venus's Atmosphere; Venus's Volcanoes; Venus's Surface Features.

Planetology: Comparative

Category: Planets and Planetology

Spacecraft have obtained detailed photographic, magnetic, radar, and chemical data from the planets Mercury, Venus, Mars, Jupiter, Saturn, Uranus, and Neptune as well as from numerous natural satellites, and even from some asteroids and comets. Data have been used in the preparation of models describing the structure and geological history of planetary and minor bodies throughout the solar system.

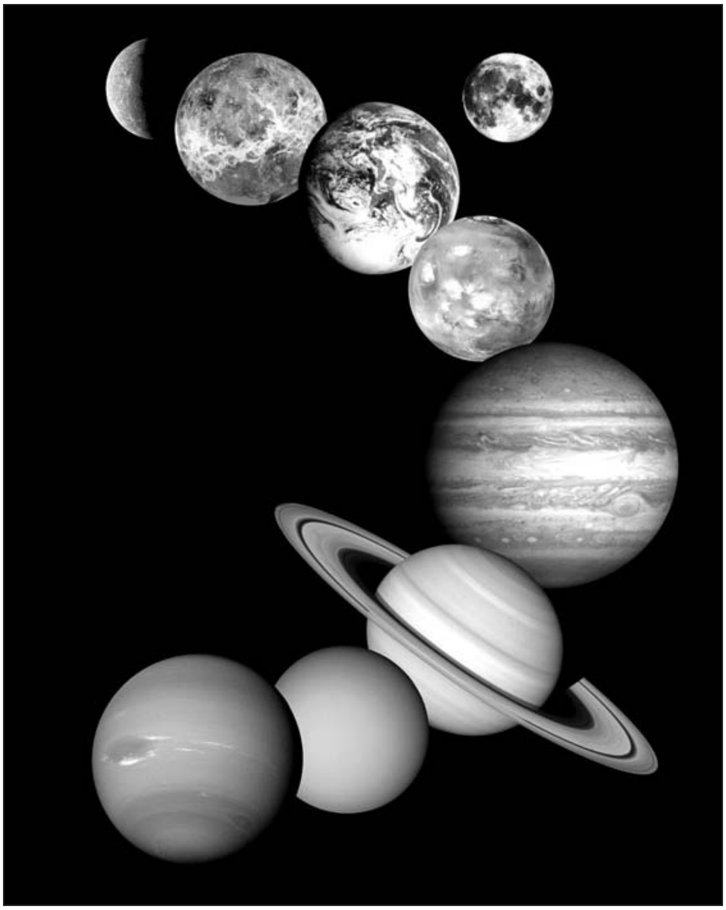
OVERVIEW

Comparative planetology is the study of the broad physical and chemical processes that operate in and on planets over time. It looks for patterns in the similarities and differences displayed by the planets and seeks to provide explanations for them in terms of planetary origins and evolution.

The first successful step in planetary exploration using robotic spacecraft was taken on August 26, 1962, when the National Aeronautics and Space Administration's (NASA's) Mariner 2 spacecraft was launched on a flyby mission to Venus. Mariner 4 was sent to Mars on November 28, 1964. Mariner, Pioneer, Pioneer Venus, Venera, Viking, Voyager, Magellan, Galileo, and Pathfinder space probes have sent back information about Mercury, Venus, Mars, Jupiter, Saturn, Uranus, and Neptune. Long after their primary missions had been completed, some of these spacecraft continued to transmit valuable information back to astronomers on Earth.

The planet Mercury had eluded detailed analysis by astronomers for centuries because of its small size and close proximity to the Sun. Mariner 10, launched in 1973 with a dual mission of studying the clouds of Venus and of photographing Mercury, made three passes by Mercury and was able to photograph about 45 percent of the surface of the planet. Mariner 10 was thus the first spacecraft to take scientific equipment to Mercury. Photographs of Mercury revealed a heavily cratered surface very much like that of Earth's moon. Naturally there are differences between the Moon and Mercury. Since the number of craters per square kilometer varies by as much as a factor of ten, it is believed that some craters may have been covered by a volcanic process. Still, Mercury's surface shows evidence of less volcanic activity than the Moon's. Mercury's largest impact basin, Caloris, has a diameter of 1,300 kilometers. It is believed to have been formed when a large asteroid struck the planet. Photographs show that the shock wave from this collision penetrated the planet and altered the surface on the opposite side, an example of antipodal focusing of seismic energy by the planet's core. Compression scarps (cliffs), which can be as much as 3 kilometers tall and hundreds of kilometers long, were also found on Mercury. They are younger than the craters and are thought to have formed as a result of some internal process such as the cooling of the core of the planet. A magnetic field with a strength of about 1 percent of Earth's magnetic field and a very diffuse atmosphere containing mostly helium were found. Surface temperature ranges from 90 to 948 kelvins.

Venus has been a difficult planet to study from Earth because of its dense atmosphere. Russian Venera probes found a surface temperature of 748 kelvins and an atmospheric pressure of 95 atmospheres. (One atmosphere is the pressure exerted by Earth's atmosphere at sea level.) Photographs of the Venusian surface revealed some areas with smooth plains, while other areas have a rocky terrain. Radar mapping first by the American Pioneer Venus probes and later, in higher resolution and with fuller coverage, by the Magellan orbiter, shows a surface that is 70 percent gently rolling plains, 20 percent lowlands, and 10 percent highlands. The Ishtar Terra highland area of Venus is larger than the United States and includes the mountain Maxwell Montes, which stands about 11.3 kilometers tall. Alpha Regio and Beta Regio are mountainous regions that may contain shield volcanoes. A Russian Vega probe observed lightning discharges in these regions, which could mean that the volcanoes are still active. Lowland areas have the appearance of a cracked slab of lava or cemented volcanic ash. The rocks in the plains are probably granitic rock or potassium-rich basalt.



The solar system planets, in a composite of images captured by missions from Mariner 10 to Cassini. (NASA/JPL)

At the relatively low altitude of 26 kilometers, the Venusian atmosphere is clear, with the temperature dropping to 583 kelvins and the pressure to 20 atmospheres. A thick cloud layer, which is about 80 percent liquid sulfuric acid in the upper portion, exists from 26 kilometers to 60 kilometers above Venus's surface. A sulfuric acid haze exists from 60 kilometers to 80 kilometers altitude. The overall atmosphere is made up of 96 percent carbon dioxide, 3.4 percent nitrogen, and trace amounts of several other gases, and reflects 76 percent of the light striking it. No planetary magnetic field was found.

Following the Pioneer Venus probes after more than a decade's hiatus, NASA returned to Venus with the Magellan orbiter. The primary objective of Magellan was to obtain a high-resolution map of nearly the entire globe of Venus using a synthetic aperture radar system. Magellan's mission was extended to permit more site-specific investigations. In all, Magellan produced a map of Venus's surface that in resolution and coverage exceeded any available maps of Earth's surface at that time.

Spacecraft such as Mariner, Viking, Pathfinder, and the Mars Exploration Rovers Spirit and Opportunity found that the surface of Mars contains craters, large plains marked by great sand dune areas, chaotic terrain characterized by irregular ridges and depressions, and many volcanoes. The largest volcano, Olympus Mons, stands 27 kilometers tall, has a base diameter of 600 kilometers, and is 64 kilometers across its summit. Orbiting probes uncovered unmistakable signs of catastrophic floods. Impact craters on Mars are concentrated in the southern hemisphere. The northern hemisphere, which has been smoothed by lava flows, has fewer craters, and their features tend to be sharper, indicating that they may be younger than those found in the southern hemisphere. Many of the craters on Mars show evidence of significant erosion.

Comparative Data on the Planets of the Solar System

Parameter	*Mercury*	*Venus*	***Earth***	*Mars*	*Jupiter*
Mass (10^{24} kg)	0.3302	4.8685	**5.9742**	0.64185	1,898.6
Volume (10^{10} km^3)	6.083	92.843	**108.321**	16.318	143,128
Equatorial radius (km)	2,439.72	6,051.8	**6,378.1**	3,396	71,492
Ellipticity (oblateness)	0.0000	0.000	**0.00335**	0.00648	0.06487
Mean density (kg/m^3)	5,427	5,243	**5,515**	3,933	1,326
Surface gravity (m/s^2)	3.70	8.87	**9.80**	3.71	24.79
Surface temperature (Celsius)	–170 to +390	+450 to +480	**–88 to +48**	–128 to +24	–140
Satellites	0	0	**1**	2	14
Mean distance from Sun millions of km (miles)	58 (36)	108 (67)	**150 (93)**	228 (141)	779 (483)
Rotational period (hrs)*	1,407.6	–5,832.5	**23.93**	24.63	9.9250
Orbital period	88 days	224.7 days	**365.25 days**	687 days	11.86 yrs

* Retrograde rotational periods are preceded by a minus sign.

Mars Exploration Rover studies of rocks in situ revealed evidence of sedimentary processes requiring the presence of water.

Mars's seasonal polar caps, made of carbon dioxide, extend well down into the hemisphere, experiencing winter, but shrink and retreat quickly in early summer. Residual polar caps remain throughout the year, although their size does vary. The southern cap is made of carbon dioxide only. The northern polar cap is larger than the southern cap, has a wider temperature variation, and contains mostly water ice. It may be one of the main storehouses for water on Mars. The Mars Polar Lander was designed to investigate that, but it crashed. In 2008, the Mars Phoenix landed on the northern polar region to continue that search for water; by sampling and analyzing the subsurface material, it provided direct evidence that water ice was present in significant amounts.

Mars's atmosphere is 95.3 percent carbon dioxide, 2.7 percent nitrogen, and 1.6 percent argon, with a total pressure of 0.01 atmosphere. The atmospheric temperature varies from 243 down to 173 kelvins. Sublimation of the polar ice caps causes the pressure to vary about 20 percent from season to season. Fog forms in low areas in the early morning. Clouds have been seen around some of the volcanoes. Winds with speeds of at least 150 kilometers per hour pick up surface dust and cause global dust storms. It can take as long as six months for all the dust to settle out from one of these storms. Several spacecraft in orbit at Mars have had to wait for months until the dust cleared in order for their cameras to resume imaging surface features. Long-lived orbital spacecraft have taken images of certain features at widely spaced intervals, showing evidence of wind erosion having altered the surface. Other features have shown evidence of water slumping of crater walls in recent times; that interpretation remains under consideration, however.

The atmosphere of Jupiter is thought to be about 1,000 kilometers thick, with a gaseous composition of 75 percent hydrogen, 24 percent helium, and 1 percent other gases. Pressure at the base of such an atmosphere would be about 100 atmospheres, and the temperature would be about 813 kelvins. The temperature at the top of the atmosphere is only about 113 kelvins. Colored bands, termed zones and belts, are visible in the atmosphere. Zones are yellow-white and represent high-pressure areas where warm currents are rising. Belts are brown, red, or blue-green and represent low-pressure areas where colder gases are sinking. Colors have their source in the interaction of chemical compounds in the atmosphere with sunlight. Very strong wind currents flow in opposite directions where the belts and zones touch. Bands are sta-

Comparative Data on the Planets of the Solar System *(continued)*

Parameter	*Saturn*	*Uranus*	*Neptune*	*Pluto* (dwarf planet)
Mass (10^{24} kg)	568.46	86.832	1,102.43	0.0125
Volume (10^{10} km^3)	82.713	6,833	6,254	0.715
Equatorial radius (km)	60,268	25,559	24,764	1,195
Ellipticity (oblateness)	0.09796	0.02293	0.01708	0.0000
Mean density (kg/m^3)	687	1,270	1,638	1,750
Surface gravity (m/s^2)	8.96	8.69	11.00	0.58
Surface temperature (Celsius)	–160	–180	–200	–238
Satellites	11	5	2	1
Mean distance from Sun millions of km (miles)	1,426 (884)	2,869 (1,779)	4,494 (2,786)	5,899 (3,657)
Rotational period (hrs)*	10.656	–17.24	16.11	–153.3
Orbital period	29.46 yrs	84.01 yrs	164.80 yrs	247.70 yrs

Source: National Space Science Data Center, NASA/Goddard Space Flight Center.

ble and have not changed their positions for the past one hundred years.

Jupiter's Great Red Spot has been its most prominent feature for over 350 years. It is about 26,000 kilometers from east to west and 14,000 kilometers from north to south. This large cyclonic storm wanders in an east-west fashion and may be stable enough to last for many more centuries. Voyager investigations were followed by the Galileo orbiter and its atmospheric probe, which entered the atmosphere at a point where it found relatively little water vapor.

Jupiter's magnetic field, which is at least ten times stronger than Earth's, produces a radiation belt that is strong enough to kill a human quickly. The radiation belt almost ruined some transistors in the Pioneer probes. Jupiter's radio emissions come from charged particles trapped in the magnetic field. Voyager 1 found a ring system whose main ring is 6,000 kilometers wide and 30 kilometers thick. A thin sheet of material extends to the surface of the planet.

Jupiter has sixty known moons. Fourteen were discovered by Earth-based astronomers, while two were found in Voyager 1 photographs. Others were found by Voyager 2, the Hubble Space Telescope, and Galileo spacecraft. Active volcanoes were found on the moon Io, Jupiter's innermost moon. An icy crust on Europa is believed to cover an ocean of liquid water; evidence of crustal movement upon such a water layer was found in 2008. That provided strong evidence for internal heating to drive large-scale movements of the icy crust.

Saturn is the second of the giant planets visited by spacecraft. Its atmospheric structure is much like that proposed for Jupiter, but its composition is more like the Sun's, with only 11 percent helium. Belts and zones seen on Jupiter are also visible on Saturn, but their colors are not as intense. The outer layer is predominantly hydrogen. The temperature at the bottom of this layer is 70 kelvins.

Saturn has the most highly developed ring system in the solar system. The ring system has a width of 153,000 kilometers and a thickness of 2 kilometers. There are nine distinct rings, labeled A through G. (Identification of portions of the planet's ring system retains naming schemes from the early days of telescopic observations, before the full complexity of the rings was seen; as a result, six parts of the overall ring system are identified by capital letters, whereas the rest are given names. Unfortunately, letters and names do not necessarily provide information as to distance from the outer atmosphere; for example, the C and B rings are outside the D ring but inside the A ring, and the F and G rings are outside the A ring.) The rings are very complex, made up of an extremely large number of ringlets, some of which are only about two kilometers wide. Shepherding moons orbit around the edge of some rings, and their gravity functions to maintain the sharp edge on the rings. The B ring shows dark features that resemble spokes in a wheel. The particle size in the rings varies from a few thousandths of a centimeter to about ten meters. The spokes rotate as if solid and appear to be particles electrostatically raised above the ring plane.

Saturn has at least sixty satellites. The Cassini spacecraft provided images of many of those satellites detected after the Voyager era. Saturn's only large satellite, Titan, has received a great deal of attention, since it is able to retain a thick atmosphere of nitrogen, methane, and other hydrocarbons. Its surface is obscured due to the thickness of that atmosphere. For that reason, the Cassini orbiter carried a European Space Agency probe named Huygens that was detached from the main spacecraft in order to land on the surface of Titan. Huygens provided evidence of liquid hydrocarbons at its touchdown site existing at cryogenic temperatures. Cassini then was able to image ancient shorelines and prove the existence of lakes of liquid methane across the surface of Titan.

In January, 1986, Voyager 2 passed by Uranus en route to Neptune. Uranus's rotational axis is tilted 82° from the plane in which it orbits. Its rotational direction is retrograde. Voyager 2 data established the rotational period of the planet to be 17 hours, 14 minutes. The greenish atmosphere of Uranus is unusually free of clouds. The primary components of the atmosphere are hydrogen (84 percent), helium (14 percent), and methane (2 percent). The temperature of the atmosphere where the pressure is 1 atmosphere is about 73 kelvins. Voyager 2 observed a tenuous haze around Uranus's rotational pole. This haze is probably formed by the steady irradiation of the planet's upper atmosphere by solar ultraviolet light. Uranus is a weak emitter of thermal radiation from deep within the atmosphere. The planet has been found to be warmer than thought, which implies a greater transparency of the atmosphere than models had predicted. The magnetic field of Uranus is inclined at a 60° angle to the axis of rotation. (Earth's rotational axis and magnetic field are roughly parallel by comparison.)

Voyager 2 found a large spot on Neptune's southern hemisphere in 1989 similar to Jupiter's Great Red Spot. The probe also confirmed the presence of three thin, faint rings around the planet and a magnetosphere. The atmosphere is cold, about 53 kelvins, and its soft, blue tint comes from the presence of methane in the upper atmosphere. Neptune has thirteen known moons, the largest of which, Triton, is covered with methane and nitrogen ices.

The New Horizons spacecraft was launched in January, 2006, to fly by the Pluto-Charon system and thereby complete the initial reconnaissance of all major systems of the solar system. New Horizons was launched when Pluto was still classified as a planet. Later that same year, a new identification system adopted by the International Astronomical Union (IAU) demoted Pluto to the status of a dwarf planet. In June, 2008, the IAU again redefined Pluto, this time as a plutoid, or plutino. Regardless of whether Pluto is a full-fledged planet or a plutoid, New Horizons will provide the first in-depth investigations and closeup images of Pluto and its nearly similar sized satellite Charon sometime in the second decade of the twenty-first century.

KNOWLEDGE GAINED

Spacecraft data concerning the atmospheric composition and structure of individual planets have provided significant insight into the solar system as a whole. Mercury's small size and high temperature made it an unlikely candidate for having any measurable atmosphere, yet Mariner 10 found a tenuous atmosphere on the planet. This condition probably arises from the solar wind that bathes Mercury. Venus has a high surface temperature but significantly more mass than Mercury and has been able to retain its atmosphere effectively. Venus's high temperature prevents the buildup of any significant quantity of water, so that carbon dioxide remains in the atmosphere rather than forming carbonates as it can on water-rich Earth. Mars perhaps once had a much denser atmosphere, with large quantities of liquid water—possibly enough to cover the planet to an average depth of 10 meters. Channels on the surface point to large amounts of flowing liquid. As a result of low temperature and low surface gravity, most of Mars's atmosphere has been lost. Perhaps water ice is still trapped below the surface or in the north polar ice cap. Confirming that was the primary objective of the Mars Phoenix mission in 2008, and early results from the lander strongly suggested that white material just underneath the soil was indeed water ice and neither salts nor dry ice. Mars Phoenix was outfitted with a Thermal Evolution and Gas Analyzer (TEGA). Before the end of 2008, TEGA obtained evidence of the presence of water vapor after heating soil samples that were carefully placed within its ovens by a robotic arm equipped with a scoop.

Since the giant planets Jupiter, Saturn, and Uranus, and Neptune have much more massive cores and are much colder than the four inner planets, they can retain light gases such as hydrogen and helium effectively. Differences exist among these four, however, because the core size differs from planet to planet.

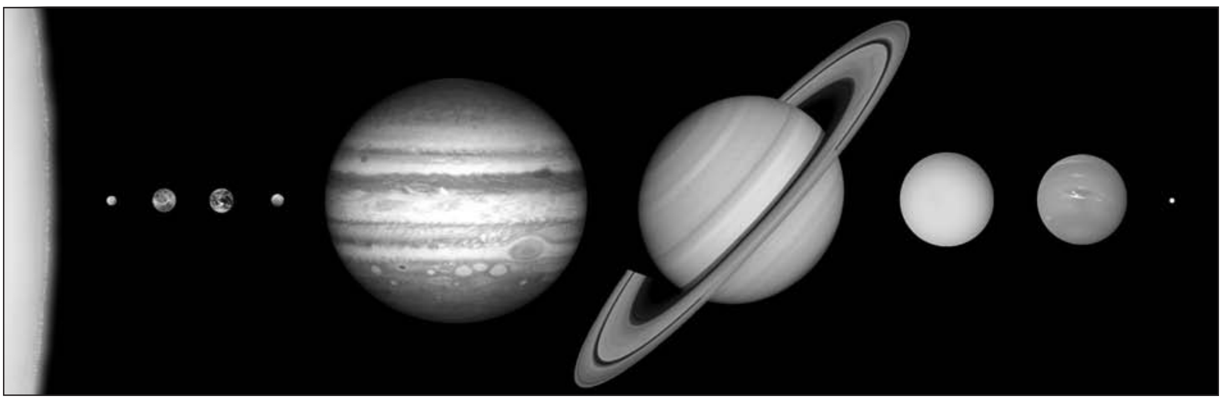
The terrestrial planets, Mercury, Venus, Earth, and Mars, show many similar surface features—craters, volcanoes, and mountains. Only Earth has shown activity of its volcanoes, but discharges of lightning around the volcanic mountains on Venus suggests that they may be active also. Volcanism has also been found on Jupiter's satellite Io, Saturn's satellite Enceladus, and Neptune's satellite Triton.

The giant planets all have ring systems, although each system is different from the other three. Saturn's rings are extremely complex, with small divisions between the rings. Uranus has a set of narrow ribbons separated by large spaces, while Neptune has only partially complete ring arcs. Jupiter has a three-component ring system. The innermost portion is called the Halo Ring. Further out is the Main Ring, and that is followed by the wispy Gossamer Rings. High-resolution images from Galileo at Jupiter, Cassini at Saturn, and Voyager 2 at Uranus and Neptune greatly added to the storehouse of knowledge about diversity in ring system dynamics.

One great hope in the exploration of Mars was that some life-form would be discovered. Experiments performed by the Viking landers provided no definitive results. Many astronomers believe that Mars's environment is much too harsh presently to support life as it would exist on Earth. Any primitive non-Earth-like forms of life might be difficult to detect. Life, primitive or otherwise, may also be possible on Europa, Enceladus, or Titan. Few scientists expect to find organisms on the latter two satellites, but some hold out hope that some degree of organized life-forms may be swimming in Europa's ocean under the satellite's icy crust. Until a Europa lander equipped with a subterranean probe can be sent to this satellite, however, that remains only wishful speculation on the part of exobiologists.

CONTEXT

Fascination with outer space is evident when one examines the popularity of space-based science-fiction books, films, and television pro-



Images of the planets (including Pluto, now classified as a dwarf planet) showing their relative sizes. The Sun's crescent appears on the left. (Lunar and Planetary Institute) grams, and when one keeps track of the number of Internet hits on NASA Web sites during high-profile missions like Mars Exploration Rover landings on the Red Planet or space shuttle flights to refurbish the Hubble Space Telescope. Solar system exploration programs are scientific attempts to satisfy human curiosity about space. One fundamental purpose for planetary exploration is to seek a better understanding of the history and perhaps the origin of the solar system. While current models meet some of the criteria, many questions remain. Examination of planetary atmospheres, magnetic fields, ring systems, satellites, and surfaces allows models to be improved and planetary history to be more accurately recorded. For example, the Jupiter and Saturn systems are large enough for them and their satellites to constitute small-scale solar systems. Study of such smaller systems could reveal significant details about the solar system as a whole.

Humanity also has a desire to know whether life exists in any place other than Earth. Are we alone? Is the vastness of the universe just for us, or is it teeming with life? Chances of detecting life in another star system are remote, even if it does exist. The search on the planets of the solar system is much more easily accomplished. In the late twentieth century, both the United States and the former Soviet Union planned uncrewed missions to Mars that would include orbiters, landers, balloons, surface-roving vehicles, and a round-trip mission to return soil samples to Earth. Many of those ambitious plans were delayed considerably, but early in the new millennium an armada of robotic spacecraft orbited around the Red Planet and a number of landers were on the surface searching for evidence of water.

Dennis R. Flentge

FURTHER READING

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. Cambridge, England: Cambridge University Press, 2004. A comprehensive work about the biggest planet in the solar system, comprising a series of articles by experts in theirs field of study. Excellent repository of photography, diagrams, and figures about the Jupiter system and the various spacecraft missions that have unveiled its secrets.

Beattie, Donald A. *Taking Science to the Moon: Lunar Experiments and the Apollo Program*. Baltimore: Johns Hopkins University Press, 2003. Explains the science gleaned from the Apollo lunar landings, including the Apollo Lunar Surface Science Experiment Packages (ALSEPs) and their results.

Briggs, G. A., and F. W. Taylor. *The Cambridge Photographic Atlas of the Planets*. New York: Cambridge University Press, 1982. A collection of the best photographs taken by space probes from the United States and the Soviet Union. In addition to the captions accompanying the photos, a discussion of the important features of each planet and its satellites is provided.

Greenberg, Richard. *Europa the Ocean Moon: Search for an Alien Biosphere*. New York: Springer, 2005. A complete description of current knowledge of Europa through the post-Galileo spacecraft era. Discusses the astrobiological implications of an ocean underneath Europa's icy crust. Well illustrated and readable by both astronomy enthusiasts and college students.

Grinspoon, David Harry. *Venus Revealed: A New Look Below the Clouds of Our Mysterious Twin Planet*. New York: Basic Books, 1998. A thorough examination of the geology of Venus that incorporates Magellan mapping and other data. Explains the Venusian greenhouse effect. A must for the planetary science enthusiast who wants an integrated approach to science and history. Includes speculation about Venus's past.

Harland, David M. *Cassini at Saturn: Huygens Results*. New York: Springer, 2007. Essentially a complete collection of NASA releases from the start of Cassini flight operations through the majority of Cassini's seventy orbits of its primary mission. Provides a thorough explanation of the entire Cassini program, including the Huygens probe's landing on Saturn's largest satellite. Cassini's primary mission concluded a year after this book was published. Technical but accessible to a wide audience.

_______. *Water and the Search for Life on Mars*. New York: Springer Praxis, 2005. A historical review of telescope and spacecraft observations of the Red Planet up through the Spirit and Opportunity rovers. Covers all aspects of Mars exploration but focuses on the search for water, believed to be the most necessary ingredient for life.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Particularly strong in its presentation of Earth-Moon science. Takes a comparative planetology approach rather than providing individual chapters on each planet. Examines atmospheres, magnetospheres, satellites, and interiors of the solar system's planets.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs.

Lovett, Laura, Joan Harvath, and Jeff Cuzzi. *Saturn: A New View*. New York: Harry N. Abrams, 2006. A coffee-table book with about 150 of the best images returned by the Cassini mission to Saturn. Covers the planet, its many satellites, and the complex ring systems.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. A discussion of data from each of the planets accompanied by a large number of photographs and line drawings. Although intended as a college-level astronomy textbook, it provides good reading for anyone with an interest in the solar system.

Squyres, Steve. *Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet*. New York: Hyperion, 2006. Written by the principal investigator for the Mars Exploration Rovers Spirit and Opportunity, this fascinating book provides a general audience with a behind-the-scenes look at how robotic missions to the planets are planned, funded, developed, and flown. A personal story of excitement, frustrations, a scientist's life during a mission, the satisfaction of overcoming difficulties, and the ongoing thrills of discovery.

See also: Planetology: Venus, Earth, and Mars; Terrestrial Planets.

Planetology: Venus, Earth, and Mars

Category: Planets and Planetology

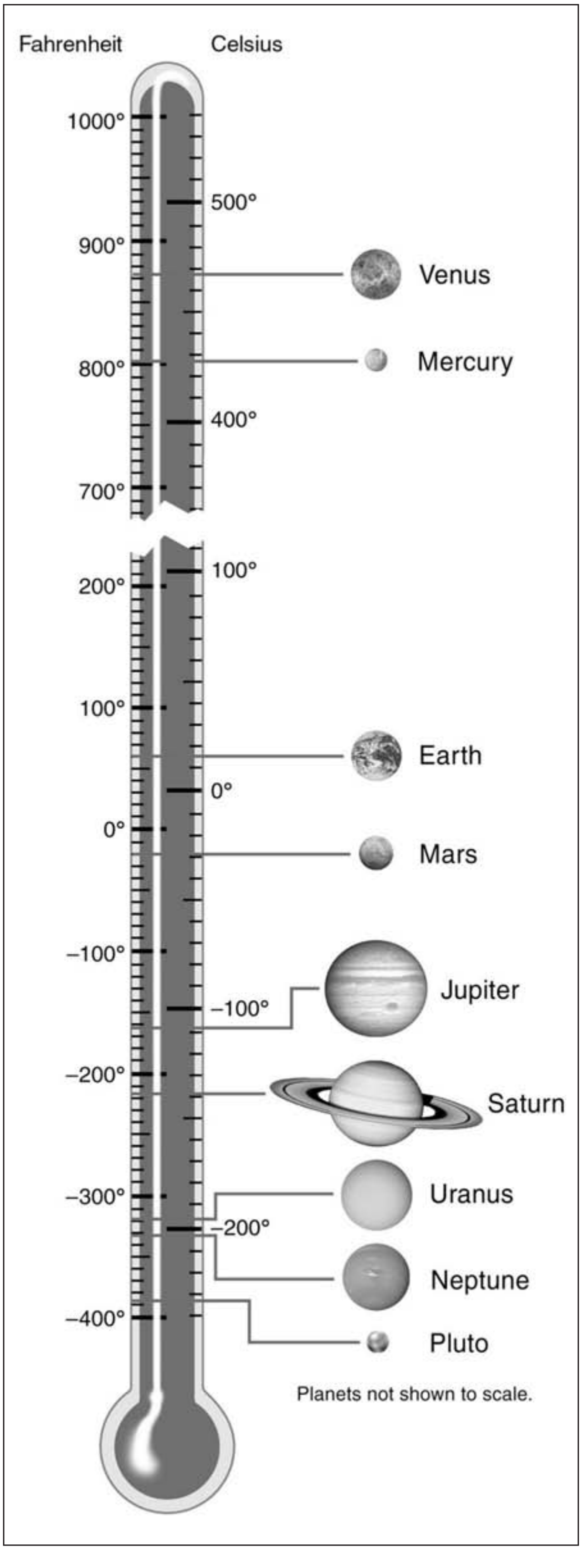
The rocky planets Venus, Earth, and Mars are similar in size, mass, and proximity to the Sun, yet they have evolved in three very different directions. Venus has a thick atmosphere of carbon dioxide and a surface temperature hot enough to melt lead. Earth's atmosphere is mostly nitrogen with only a trace amount of carbon dioxide. Mars has a very thin atmosphere of carbon dioxide and quite cold surface temperatures.

OVERVIEW

When the planets formed in the primordial solar system, they existed primarily as gases coalesced by gravity. The four inner planets, because of their proximity to the Sun, lost most of their primitive atmosphere of hydrogen and helium, retaining only a molten rocky core rich in heavier elements. As these planets cooled, a secondary atmosphere was created by gases ejected from the mantle by many active volcanoes on the geologically unstable surface. These gases included ammonia, carbon dioxide, water vapor, and nitrogen. Mercury, being the smallest planet with the greatest intensity of sunlight, quickly lost its entire atmosphere.

More massive and farther from the Sun, Venus retained its secondary atmosphere. This caused the surface to remain hot enough to prevent the water vapor from condensing into liquid. As the concentration of CO_2 increased, the heat in conjunction with ultraviolet light from the Sun dissociated the water vapor into free hydrogen and oxygen. The hydrogen dissipated into space, while the very reactive oxygen combined with rock minerals, thus disappearing from the atmosphere. Estimates made from Venus's rocks indicate that initially there was enough water vapor on Venus to have covered the surface with an ocean at least 9.14 meters deep, on average. The continued volcanic outgassing of CO_2 continued to heat the surface until equilibrium was established. When visible light strikes a planetary surface, much of the energy is converted into heat, which radiates from the surface as infrared radiation. Carbon dioxide will absorb outgoing infrared radiation and reemit it uniformly in all directions, thus heating the surface in proportion to the amount of CO_2 present in the atmosphere. Because Venus's atmosphere is 95 percent carbon dioxide, the surface temperature is even hotter than the surface of the planet Mercury, due to what is referred to as a "runaway greenhouse effect." Venus therefore is now a stifling inferno smothered by a thick carbon dioxide (CO_2) atmosphere. The surface temperature is about 750 kelvins, distributed uniformly across the planet due to atmospheric carbon dioxide.

Because Earth is 43 percent farther from the Sun, the solar intensity is only half that at Venus, while the gravitational attraction is slightly greater than Venus's. Consequently, water vapor was retained and able to condense as the planet cooled, thereby forming the oceans. Carbon dioxide readily dissolves in warm water. It was progressively removed from the atmosphere, eventually forming carbonate rocks. Relatively early in Earth's history primitive life formed in the oceans, and evolved into small shelled sea creatures. These primitive animals formed their shells from the dissolved carbon dioxide and calcium, thus further removing CO_2 from the water and leaving shell deposits which eventually agglomerated into calcium carbonate ($CaCO_3$). As CO_2 was removed from the water, more could enter from the atmosphere until almost all was gone. Analysis shows that the carbonate rocks on Earth's crust contain about the same amount of CO_2 as Venus's atmosphere. Although oxygen was pres-



Temperatures of the planets. (Lunar and Planetary Institute) ent in Earth's primitive atmosphere, most reacted with metals, such as iron, to form oxides. Life further altered the atmosphere with photosynthesis commencing about 2.5 billion years ago, eventually reducing the concentration of atmospheric CO_2 to its present 0.04 percent and creating most of the free oxygen.

Mars is 1.5 times farther from the Sun than Earth. Being considerably smaller and less massive, Mars had less internal heating and consequently considerably less outgassing. Mars therefore has an extremely thin atmosphere, 95 percent of which is CO_2. This is an amount insufficient to create any significant greenhouse warming.

The Viking spacecraft showed that although Mars originally had a denser atmosphere and running water on the surface, most of the atmosphere leaked off into space because of the low gravity. The atmospheric pressure, less than 1 percent of Earth's sea-level pressure, is below the pressure where water can exist in the liquid state. Although a core of water ice is present in the polar caps, the ice in those caps consists primarily of solid carbon dioxide. Surface temperatures on Mars vary from a maximum of 300 kelvins at the warmest spot at the warmest moment of the warmest day of the Martian year to typical nighttime temperatures of 155 kelvins. At a location midway between the equator and the poles, the maximum daytime temperature barely exceeds the freezing point of water. Much of the remaining water is assumed to be tied up below the surface as permafrost. The original CO_2 is now either carbonate rocks, or dissipated into space. The polar caps consist of small permanent water ice caps and solid CO_2 (dry ice), which sublimates directly into its vapor form when the temperature increases during summer.

Although the atmospheres of Venus, Earth and Mars appear very different, when the elements bound up in rocks and permafrost are included, the inventories of water, carbon dioxide, and nitrogen are remarkably similar when adjusted for the differing planetary masses.

The topography of the three planets is a result of their size and evolution. Earth, the largest, is dominated by rolling seafloor plains interrupted by continents and mountain ranges where continental plates have collided. All traces of Earth's primeval crust have been destroyed by basalts that erupted to form much of the seafloor crust and by plate tectonics that broke up and recycled the original surface.

On Venus, basaltic volcanism dominated and covered most of the planet with lava flows. Only a small part of the primordial surface remained as protocontinents projecting several miles above the plains. Perhaps because of its smaller size, there was not enough internal energy to drive plate tectonic crustal motions. Some volcanoes may still be active. The scarcity of meteorite impact craters indicates that the entire crust has been replaced within the past half million years—recently, in geologic terms.

Being considerably smaller than Venus, Mars lacked sufficient tectonic energy to destroy its original cratered features. The deepest surface depressions are ancient impact craters but the highest Martian mountains are simply masses of volcanic lava surmounted by the volcanic caldera. As a rule of thumb, small worlds preserve their ancient surfaces formed by meteorite bombardment, but volcanic forces break through and resurface parts of the planet. Larger planets, on the other hand, have surfaces dominated by the internal forces of volcanism and plate tectonics.

KNOWLEDGE GAINED

It has been known since the eighteenth century that Venus has an atmosphere. Although it was first assumed to be composed of water vapor, in 1932 spectroscopic studies indicated that Venus's atmosphere was primarily CO_2. Thermal radiation measurements in the 1960's indicated a surface temperature close to 750 kelvins. This temperature was confirmed when, in 1967, the Russian probe Venera 4 crashed into the Venusian surface. In 1970 the Venera 7 successfully landed on Venus's surface and transmitted data for twenty-three minutes, verifying the surface temperature and an extremely dense CO_2 atmosphere. The American spacecraft Pioneer (1978) discovered clouds of sulfuric acid droplets positioned about 48 kilometers above surface. This mission also included an orbiter that mapped the surface features by using radar. Russian landings later sent back panoramic photographs of a haze-free surface imbued with a reddish hue from atmospheric filtered sunlight, strewn with boulders on gravel and fine, rocky soil. In 1985, two balloons dropped by Russian probes floated in the sulfuric acid clouds for forty-six hours, measuring hurricane-force winds of 240 kilometers per hour but temperatures and pressures similar to Earth's surface. When the sulfuric acid particulates reach a sufficient size, they begin to fall out of the cloud deck as sulfuric acid rain. Unlike rain on Earth, however, they never reach the ground. The rapidly increasing temperature cause them to evaporate.

Telescopic observations of Mars during the eighteenth and nineteenth centuries revealed clear seasonal changes. In summer the polar cap shrinks, while dark markings, once assumed to be vegetation, darken and grow more prominent. This assumption was disproved in 1965 when Mariner 4 returned the first close-up pictures revealing a desolate, red-colored desert. Three Russian probes reached the surface between 1971 and 1974, but all failed to return useful data. The first successful landing was by Viking 1 on July 20, 1976, followed in September by its sister craft, Viking 2. The Viking landers photographed a surface strewn with boulder-sized fragments of lava flows and meteorite impact craters in different stages of erosion. Soil analyses provided chemical evidence that Mars's atmosphere was once almost as dense as Earth's, which would have enabled water to flow. Earlier, Mariner 9 had photographed channels looking like dry riverbeds containing tributaries and sedimentary deposits, providing additional evidence that water once flowed on Mars.

Indisputable evidence of Mars's water was provided during the summer of 2008, when National Aeronautics and Space Administration's (NASA's) Mars Phoenix Lander dug into Martian soil and scooped up what appeared to be ice. When melted it proved to be water vapor. Three spacecraft landed between 1997 and 2004 to photograph and analyze the surface using robotic rovers. The Martian surface consists of two parts: a heavily cratered southern highlands and northern smooth lowlands. Evidence from the Odyssey spacecraft (2001) suggests these lowlands were once filled by an ocean of liquid water, now present in permanent ice caps and subterranean permafrost. Mars's early dense atmosphere may have been depleted by changes in the polar climate. If solar radiation striking the poles were 10 percent higher in the past, the ice caps would have been much smaller and most of the now frozen CO_2 would have existed in a gaseous state.

CONTEXT

Although Venus, Earth, and Mars are similar in mass and distance from the Sun, enough minor differences existed to compel each to have evolved along very different paths. The atmospheres of Venus and Mars have been thoroughly analyzed, and their surface features have been carefully mapped, photographed, and subjected to chemical analysis. Nevertheless, several questions remain unanswered. How much water was initially present on Venus, and what happened to it? Why was there not more plate tectonic activity during Venus's formative years? Is Venus still volcanically active? Why does Earth's atmosphere contain an abundance of nitrogen, and why is it scarce on the other planets? Why did Mars have a much warmer surface in the past, and where is the once prevalent water? Did the formation of the gigantic volcano, Olympus Mons, cause a change in the tilt of the Martian axis that instigated a planetary cooling? Is Mars still volcanically active, or has it cooled sufficiently that new volcanoes cannot form?

Obviously, there is still much to learn about the comparative planetology and evolution of these three planets. Future interplanetary spacecraft that land on Venus and Mars will continue to advance the understanding of the dynamics of their surfaces and, in particular, the manner in which subterranean features influenced their respective planetary evolution.

George R. Plitnik

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Includes an excellent summary of the latest knowledge about the terrestrial planets, particularly Venus, Earth, and Mars.

The atmosphere, the surface, and the internal structure of each of these planets is discussed in great detail, rendering the comparative planetology a straightforward exercise for the dedicated reader.

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001. A challenging and thorough text for students of planetary geology. Covers extrasolar planets and provides an in-depth contemporary explanation of solar-system formation and evolution. Best for the most serious reader with a strong science background.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that had investigated all planets with the exception of Pluto by the publication date. Takes an astrophysical approach to place our solar system in a wider context as just one member of similar systems throughout the universe.

Esposito, Larry W., Ellen R. Stofan, and Thomas E. Cravens, eds. *Exploring Venus as a Terrestrial Planet*. New York: American Geophysical Union, 2007. This is a collection of articles covering all major areas of Venus research. Technical.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. This authoritative and regularly updated text considers all the major planetary objects in the solar system. The material is presented by grouping objects under unifying principles, thus elucidating their similarities and their differences as well as the physical processes behind their evolution. Although most of the material is descriptive, some algebra and elementary calculus are included.

Lewis, John S. *Physics and Chemistry of the Solar System*. 2d ed. San Diego, Calif.: Academic Press, 2004. Suitable for an undergraduate course in planetary atmospheres, but accessible to the general reader with a technical background.

McSween, Harry. *Stardust to Planets*. New York: St. Martin's Griffin, 1993. A well-written book that imparts geologic and atmospheric facts about the planets of the solar system in an easy-to-grasp manner. Venus and Mars are covered in considerable detail, and the information about Earth's geology makes planetary comparisons relatively straightforward.

Sagan, Carl. *Cosmos*. New York: Wings Books, 1995. Based on the television series of the same name, this lavishly illustrated classic includes considerable information about the evolution of surface features and atmospheres of Venus and Mars.

Seeds, Michael A. *Foundations of Astronomy*. 9th ed. Belmont, Calif.: Thomson Brooks/Cole, 2007. This lavishly illustrated text commingles experimental evidence and theory to provide deep, but well-explained, elucidations of many fascinating aspects of the universe. In particular, the five chapters on comparative planetology contain a plethora of data, as well as beautiful pictures, from various spacecraft encounters with these planets since the 1970's.

See also: Planetology: Comparative; Terrestrial Planets.

Pluto and Charon

Categories: Natural Planetary Satellites; Planets and Planetology; Small Bodies

Pluto became the ninth planet in the solar system when discovered in 1930. Then, in 2006, the International Astronomical Union demoted Pluto to the status of a dwarf planet. Pluto and its satellite Charon constitute a dual-object system located far from the Sun. These bodies are different in size and composition from any of the planets of the solar system, more closely resembling the icy satellites of Neptune.

OVERVIEW

Pluto was discovered by American astronomer Clyde Tombaugh in 1930; its satellite Charon was detected in 1978 by James W.

Christy of the U.S. Naval Observatory. In the United States, “Charon” is pronounced SHAR-uhn, reminiscent of the discoverer’s wife, Charlene; in the rest of the world, the pronunciation KAR-uhn is usually preferred.

Less is known about Pluto than about any of the other planets; it is the only planet that has not been visited by a satellite from the Earth. Earth-based telescopes cannot provide much information about Pluto and Charon, as they are too far away for surface details to appear even in the largest telescopes. Better images of Pluto were obtained using the Hubble Space Telescope, which in May, 2005, detected two new small satellites of Pluto. Hubble images also provided the first indications of features on the surface of Pluto. Almost a dozen “provinces”—portions of Pluto with different albedos—were discovered. Pluto also was observed to have a northern polar cap, several dark spots, and a bright linear feature. The resolution of these Hubble images precluded a precise determination of the nature of these intriguing features. However, scientists speculated that the images were hinting at basins or craters of significant size. Identification of these features would have to wait for the New Horizons spacecraft to fly through the Pluto-Charon system in 2015.

Pluto is the smallest planet in the solar system, smaller even than Earth’s moon; it is usually the outermost planet. Pluto takes 247.7 Earth years to orbit the Sun and rotates on its axis once every 6.39 Earth days. The orbit of Pluto is more eccentric than that of any other planet. Pluto’s orbital eccentricity is so large that Pluto is sometimes closer to the Sun than Neptune: That was the situation between January 21, 1979, and March 14, 1999, when Pluto’s orbit again took it farther from the Sun than Neptune. Pluto will remain beyond the orbit of Neptune until 2226.

Distances between objects in the solar system are usually measured in astronomical units (AU). One astronomical unit is the average distance between the Earth and the Sun, or 150 million kilometers. Pluto can be as close to the Sun as 29.64 AU and as far away as 49.24 AU. At Pluto’s distance, the Sun appears as a starlike point, but a point more than one hundred times brighter than a full Moon. The amount of solar energy received by Pluto varies greatly because of the large variation in distance between

The Hubble Space Telescope captured this image of Pluto (lower left) and Charon, more than 4 billion kilometers away, in 1994. (Dr. R. Albrecht, ESA/ESO Space Telescope European Coordinating Facility; NASA) the Sun and Pluto over the course of its lengthy "year" (orbit around the Sun). This variation in solar energy is expected to cause the thickness of Pluto's atmosphere to change markedly in different parts of its orbit.

Pluto Compared with Earth

Parameter	*Pluto*	*Earth*
Mass (10^{24} kg)	0.0125	5.9742
Volume (10^{10} km^3)	0.715	108.321
Equatorial radius (km)	1,195	6,378.1
Ellipticity (oblateness)	0.0000	0.00335
Mean density (kg/m^3)	1,750	5,515
Surface gravity (m/s^2)	0.58	9.80
Surface temperature (Celsius)	–238	–88 to +48
Satellites	3*	1
Mean distance from Sun millions of km (miles)	5,899 (3,657)	150 (93)
Rotational period (hrs)	–153.2928	23.93
Orbital period (days)	90,588	365.25

*In May, 2005, the discovery of two small Plutonian moons, Nix and Hydra, brought the total (with Charon) to three.

Source: National Space Science Data Center, NASA/Goddard Space Flight Center.

Charon is similar in size to Pluto. The diameter of Pluto is approximately 2,284 kilometers, and Charon's is approximately 1,192 kilometers. The average distance between Pluto and Charon is 19,700 kilometers. Because of this close proximity, along with their similar sizes, Pluto and Charon are referred to as a double planet. Charon orbits Pluto in 6.39 days—the orbital period of Charon is the same as the rotation period of Pluto. As a result, Pluto always points the same face toward Charon. In fact, an observer on the surface of Pluto would always see Charon in the same position relative to the horizon.

The orbit of Charon is not in the plane of Pluto's orbit about the Sun. Instead, the plane swept out by Charon's orbit is almost perpendicular to the plane swept out by Pluto's orbit. When the plane of Charon's orbit presents its edge to the Earth, a series of occultations and transits between Pluto and Charon occur. This series of transits and occultations results in a series of mutual eclipses being observed from the Earth. These mutual eclipses can be observed at two positions in Pluto's orbit, and so they occur every 124 years. Mutual eclipses last for about 6 years. A series of mutual eclipses that began in 1985 made possible the measurement of the sizes of Pluto and Charon reported above.

The surface temperature of Pluto is somewhat uncertain because the fraction of sunlight that it reflects (known as the albedo) is uncertain. Pluto's surface temperature ranges from 45 to 60 kelvins. The uncertainty arises because the surface composition of Pluto and the extent of its atmosphere are uncertain. The surface temperature of Charon is estimated to be between 8° and 10° warmer than Pluto's.

The density of the Pluto-Charon system has been determined to be of the order of 1,800 kilograms per cubic meter. This density, almost twice that of water, indicates that Pluto and Charon are composed of a variety of ices and that as much as half of their mass could be made up of rocky material. The surface of Pluto has, in fact, been determined to contain methane ice. It is thought that, rather than the methane being uniformly distributed over the surface, there are two large polar ice caps made of methane and a thin, warmer equatorial region, where the methane has become depleted, leaving water ice.

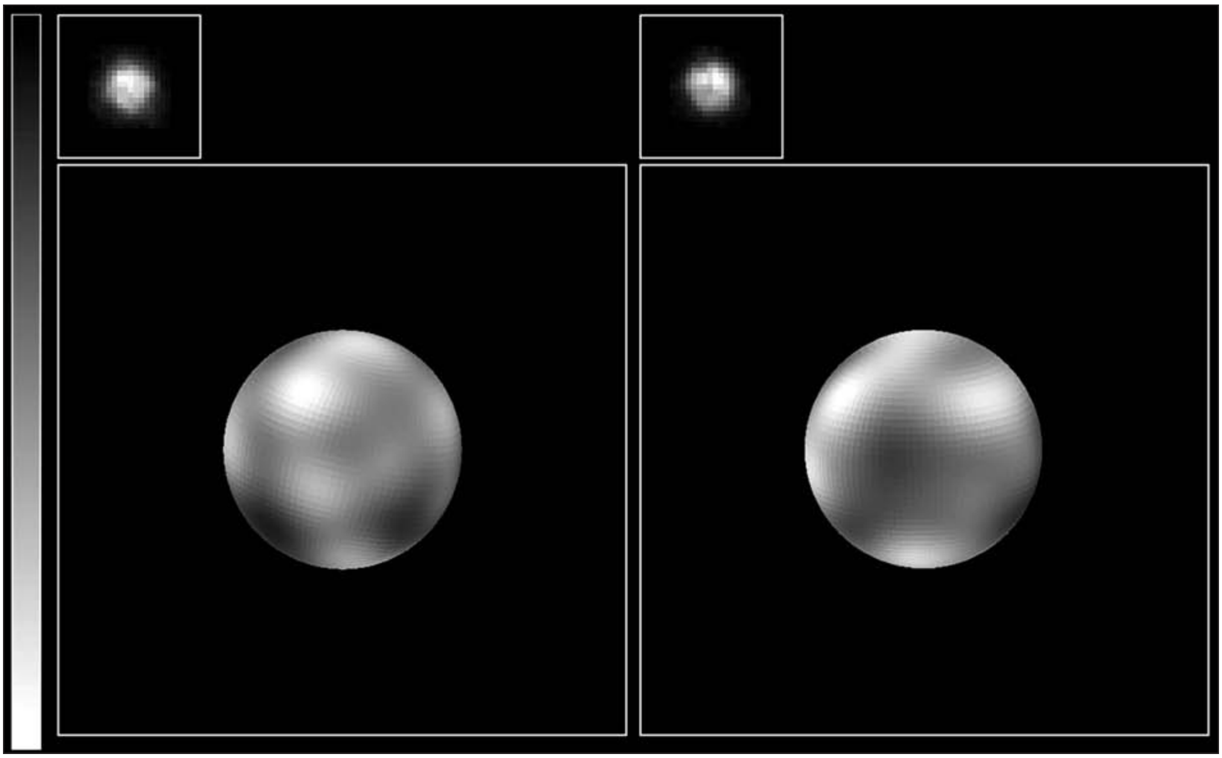
Pluto is too small to trap a permanent atmosphere, but in the late 1980's, a thin atmosphere of methane was detected. Scientists believe that the atmosphere of Pluto was at its thickest during this period because Pluto was near its perihelion passage. At perihelion, its closest approach to the Sun, Pluto receives more energy from the Sun than it does during other parts of its orbit; in other words, it is heated more strongly. Methane is frozen under the conditions prevailing on Pluto. When it is heated sufficiently, it will form a gas directly from the solid without first forming a liquid, a process is called sublimation. It has been theorized that the atmosphere detected on Pluto may result from sublimation of methane from the equatorial region of its surface to form an atmosphere. Pluto will have to be observed through its entire orbit before it is known if it has an atmosphere throughout that orbit, and if it does, how the atmosphere's thickness specifically varies. It has been suggested that when Pluto is close to aphelion, only the side of Pluto facing the Sun would be warm enough to maintain a methane atmosphere. The atmosphere on the far side of Pluto would precipitate on its surface as frost.

The surface of Charon has been determined to be covered with water ice; no frozen methane has been detected. It is expected, however, that the interior of Charon contains methane. The composition of Charon is similar to that of some of the satellites of the Jovian planets. In fact, the surface of Charon appears to be almost identical in composition to that of Miranda, one of the satellites of Uranus. Charon is not expected to trap an atmosphere, even temporarily. It is difficult to make an exact determination, but an upper limit of no more than one-twentieth of the thickness of Pluto's atmosphere has been determined. Pluto's equatorial region is depleted in methane and thought to have the same composition as its similarly methane-depleted satellite, Charon.

METHODS OF STUDY

Most of the information currently available about the dual planet has been derived from the electronic recording of telescopic images of Pluto and then computer processing of these images. The rotation period of Pluto was measured by noting that the brightness of Pluto varies periodically with the rotation period of the planet. The brightness varies because the surface distribution of methane ice and water ice is not uniform, and different ices reflect different amounts of light.

Charon was discovered while James Christy examined some electronic images of Pluto. He noticed a bump on the edge of Pluto that appeared to move; this "bump" was Charon. No



The 1994 Hubble images of Pluto (small images, upper-left insets), have been enhanced in the larger images to reveal surface and atmospheric features. (NASA) ground-based telescope was able to separate Pluto and Charon into well-resolved images, however, because of the Earth's atmosphere.

The atmospheres of Pluto and Charon have been studied by two different methods: occultations and spectroscopy. An occultation occurs when the light from an astronomical object is extinguished by another celestial object, such as when Pluto passes in front of a star. The observation of occultations is the standard technique used to determine whether a planet has an atmosphere or rings. If the planet has no atmosphere, it is possible to observe the star with undiminished brightness until the disk of the planet crosses it. It then disappears completely and reappears with its usual brightness. If a planet has an atmosphere, light from the star dims gradually as starlight passes through the atmosphere of the planet. When it reappears, it is faint and brightens as the planet moves farther away from the star. The atmosphere of Pluto was first detected in this manner.

Some Facts About Charon

Mean distance from Pluto (km)	19,600
Sidereal orbit period (days)	6.38725
Sidereal rotation period (days)	6.38725
Orbital inclination to Pluto (degrees)	0.0
Orbital eccentricity	0.0
Equatorial radius (km)	593
Mass (10^{21} kg)	1.62
Mean density (kg/m^3)	1,850
Surface gravity (m/s^2)	0.31
Escape velocity (km/s)	0.60
Albedo	0.38
Apparent visual magnitude	16.8

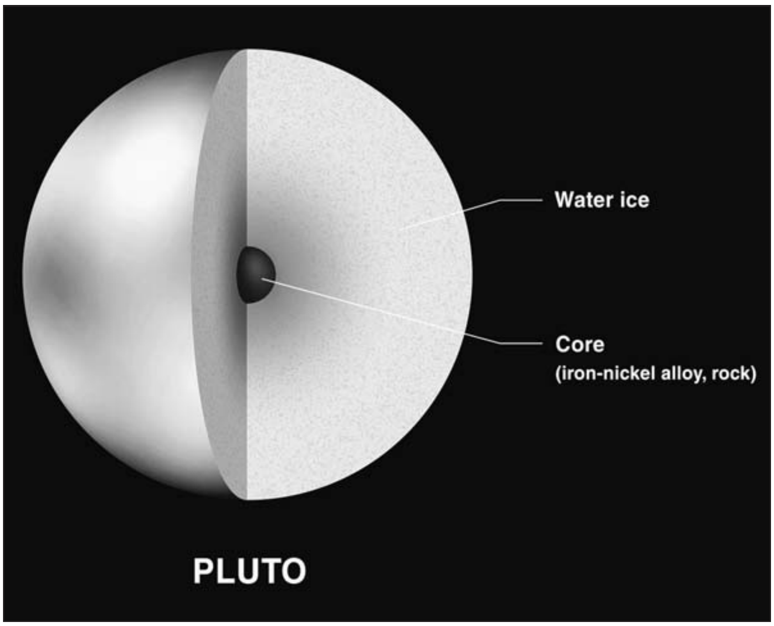
Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.

Spectroscopy involves analysis of light reflected by Pluto. Different wavelengths are reflected by different degrees, and some are completely absent from the reflected light. The spectrum of reflected light can be used to identify chemical elements and compounds present on the surface of a planet and in its atmosphere. This procedure works because each element or compound produces a unique spectrum that can be measured in the laboratory. The infrared spectrum of Pluto has also been probed to add to the information. The main problem encountered in the Pluto-Charon system is that normally the spectra of Pluto and Charon are obtained simultaneously. Mutual eclipse events described above have enabled the spectrum of Pluto alone to be obtained when Charon is behind Pluto. This Pluto spectrum can then be subtracted from the usual combined spectra to obtain the spectrum of Charon. Using this method, scientists have been able to determine the different surface compositions of Pluto and Charon.

Occultations could also be used to measure the sizes of Pluto and Charon, but instead scientists have used the series of mutual eclipses. The rotation period of Charon about Pluto is known, so if the durations of the eclipses of Charon by Pluto are timed (and visa versa), these times can then be used to estimate the diameters of Pluto and Charon. Masses of the outer planets are usually measured by their effects on the orbits of planets closer to the Sun. This method, however, has not worked in the case of Pluto and Charon, because their combined mass is too small to have an observable effect on the next closest planet, Neptune. Fortunately, however, the discovery of Charon made it possible to determine the mass of Pluto from the orbital period of Charon. Kepler's third law of planetary motion states that the square of a planet's orbital period divided by the cube of its orbital radius is equal to a constant. The constant depends on the mass of the object orbited; hence, scientists have found that the mass of Pluto is about one five-hundredth Earth's mass. The mass of Charon has been determined from its size by assuming it has the same density as Pluto.

Prior to Pluto's demotion in 2006 to dwarf planet status, it had often been said that it was the only planet yet to be explored by spacecraft. Indeed, due to the tremendous distance between Pluto and Earth, the best way to gain information about Pluto and Charon would be an instrumented spacecraft flying close to the illusive system or even perhaps entering long-term orbit about Pluto. A number of proposals were entertained to conduct the flyby, including one that would use a gravity assist from Jupiter in order to get to Pluto before Pluto's orbit took it so far from the Sun that its atmosphere would freeze and cover the surface. A Pluto flyby was approved and funded, then canceled, followed by a new proposal that in turn failed to come to fruition. Eventually, the New Horizons spacecraft was proposed, approved, designed, and then launched on January 19, 2006. The spacecraft left Earth behind with the fastest speed ever attained by a human-made object. It passed the Moon in a mere few hours and reached Jupiter in only one year. Nevertheless, New Horizons will not arrive in the Pluto-Charon system until 2015.

New Horizons, if successful, will provide a wealth of new information about the Pluto-Charon system. Scientific instruments incorporated into the spacecraft include a long-range reconnaissance imager, a near-infrared imaging spectrometer, an ultraviolet imaging spectrometer, an electrostatic analyzer, a time-of-flight ion and electron sensor, a radio science experiment, and a dust counter. Incredibly, the total mass of these instruments is a mere 31 kilograms, and they operate on only 21 watts of electrical power.



Pluto's projected composition; surface methane ice sublimates, creating a thin methane atmosphere, when the dwarf planet is facing the Sun. (Lunar and Planetary Institute)

CONTEXT

Once considered the outermost planet, Pluto remains the most difficult to investigate. What has been learned about it indicates that it is different from all the other planets. The other four planets of the outer solar system—the "gas giants" Jupiter, Saturn, Uranus, Neptune—have low densities and are composed primarily of gases. The density of Pluto is greater, indicating the presence of some rocky material. Nevertheless, the density is lower than the densities of the terrestrial planets of the inner solar system (Mercury, Venus, Earth, and Mars). Pluto has features in common with the Galilean satellites of Jupiter and some of the satellites of Saturn, Uranus, and Neptune, but none has exactly the same makeup. Charon is in many ways similar to an asteroid. In fact, models developed for asteroids covered with water ice are applicable to Charon, and they have been used profitably in an attempt to gain a deeper understanding of Charon.

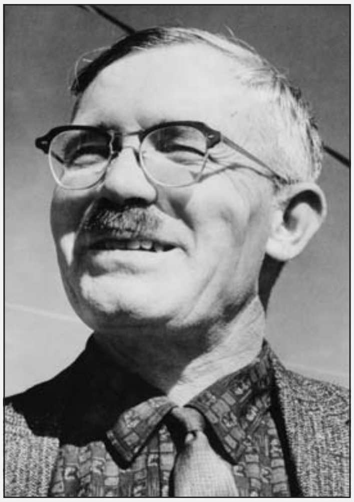
This pair of small worlds may seem insignificant in comparison with the other, larger planets of the outer solar system. However, if scientists are ever to develop a complete understanding of the origins and evolution of the solar system, they will need detailed knowledge of all its members. Pluto is usually the farthest planet from the Sun, but in many ways it has more in common with comets, meteoroids, and asteroids than it does with the other major planets. Pluto could well be one of a large number of similar objects in the outer solar system. Although there are no large objects in the solar system even ten times farther away from the Sun than is Pluto, there is a spherical swarm of comets far from the Sun, called the Oort Cloud. In developing an understanding of Pluto and Charon, scientists may be laying the groundwork for a better understanding of the Oort Cloud.

Study of the Pluto-Charon system may also answer some questions about Neptune and its satellites. One of the theories for the origin of Pluto is that it was once a satellite of Neptune. That appears less likely since scientists have learned that Pluto has its own satellite. However, an understanding of its eccentric orbit would provide a definitive answer. The mysteries of the solar system include the eccentric orbit of Neptune's satellite Nereid and the clockwise direction (as viewed from the north pole) of the orbit of another satellite, Triton, when all the other large satellites of the solar system orbit their planets in a counterclockwise direction. Both of these oddities could be explained by a collision in which Pluto broke free. However, that theory lost favor with the discovery of Quaoar, Sedna, and Eris and the realization that the Kuiper Belt is likely populated with almost countless bodies of lesser size than these three icy objects. The common belief presently held is that Pluto and Charon were created in the early Kuiper Belt by a giant impact, not unlike the impact theory for the formation of the Earth and its Moon. More research is needed to pinpoint the origin of this unusual dual-object system in the outer solar system.

It was in part for these reasons that the International Astronomical Union (IAU) moved to reclassify Pluto as a dwarf planet even though it still met many of the restated criteria for classification as a planet. That definition now requires that a body orbit about the Sun, that it is in hydrostatic equilibrium, and that it has cleared out its environment. Pluto meets only the first two of those criteria. In addition, since 2000 three objects roughly the same size as Pluto have been discovered well beyond Pluto's orbit. These bodies (Quaoar, Sedna, and Eris) are believed to be Kuiper Belt objects, and Pluto

Clyde Tombaugh: Searching for Planet X

In many ways, Clyde Tombaugh was like his astronomer hero William Herschel, who discovered Uranus unexpectedly during a routine sky survey in 1781. Both were dedicated amateur astronomers and skilled telescope makers who devoted hours to tedious observations. Tombaugh, however, was only twenty-four years old when he discovered Pluto, while Herschel was in his early forties. Furthermore, Tombaugh's yearlong search for Planet X lasted much longer than that of either Herschel or Johann Galle, who discovered Neptune in 1846 on the first night he looked for it, lying less than 1 degree from its predicted position.



(National Archives)

The search for Pluto was complicated by the fact that its orbit is highly eccentric—sometimes even passing inside the orbit of Neptune—and has a large inclination of about 17 degrees from the mean plane of the other planets. It is now known that Percival Lowell's predictions for the position of Pluto were based on faulty calculations, and its discovery within 6 degrees of the predicted location was only a coincidence. Fortunately, Tombaugh did not limit his observations to the predicted area of the sky or to the region close to the mean orbital plane of the planets.

When James Christy discovered Pluto's moon, Charon, in 1978, it was conclusively demonstrated that the mass of Pluto was far too small to cause observable deviations in the orbits of Uranus and Neptune; thus the two larger planets' orbits could not be used to predict Pluto's position. In the 1990's, several icy objects much smaller than Pluto were discovered just beyond its orbit in the Kuiper comet belt with periods of about 300 years, compared with Pluto's 248-year period.

is likely the first of the Kuiper Belt objects to have been discovered. As such, it is the model for the rest of what may be a huge class of similar objects at ever-increasing distance. Some have coined the term "plutinos" to describe icy bodies coming from the Kuiper Belt. Because of the new classification scheme adopted by the IAU in 2006, the solar system officially consists of eight planets, at least three dwarf planets (Ceres, Pluto, and Eris), and numerous satellites, asteroids, comets, and minor bodies.

In an attempt to clarify its controversial planetary classification and minimize the worldwide discontent expressed over Pluto's elimination from full-fledged planet status, an executive meeting of the International Astronomical Union (IAU) held in June, 2008, in Oslo, Norway, proposed defining a plutoid as a solar system body beyond the orbit of Neptune having enough mass to assume a nearly spherical shape but not able to clean its orbit of other material, as the eight planets have. Pluto would therefore be the first plutoid object. Eris would be the second. Under this definition, Ceres would remain a dwarf planet but could not be considered to be a plutoid, since it exists in the asteroid belt between Mars and Jupiter. For many, this proposal did little to mitigate earlier objections. This classification scheme left the solar system with only the one dwarf plantet, which until 2006 most astronomers had been quite content to consider to be the largest asteroid. More plutoids were expected to be found, located even farther out in the Kuiper Belt beyond Eris, as technology permitted their detection. In the period following the IAU's adoption of the reclassification scheme, there remained much controversy, not only among professional astronomers but also among amateur astronomers, teachers, and schoolchildren.

Whether or not Pluto is a planet, its further investigation is hindered until the New Horizons spacecraft can fly through the Pluto-Charon system in 2015. Clyde Tombaugh provided the first evidence of Pluto's existence. Some proposed that a spacecraft be sent to explore Pluto at close range, and that it be christened Tombaugh. Clyde Tombaugh passed away in 1997 before such a flyby mission was approved and launched. However, the New Horizons probe presently is on track for a 2015 encounter with Pluto and Charon. On board that spacecraft are some of Tombaugh's ashes. A flyby of Pluto would complete humanity's first exploration of what once was considered a nine-planet solar system.

Stephen R. Addison

FURTHER READING

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001. A challenging and thorough text for students of planetary geology. An excellent reference for the most serious reader with a strong science background.

Faure, Gunter, and Teresa M. Mensing. *Introduction to Planetary Science: The Geological Perspective*. New York: Springer, 2007. Designed for college students majoring in Earth sciences, this textbook applies general geological principles to bodies throughout the solar system. Excellent for learning comparative planetology.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Provides a full description of our contemporary understanding of Pluto.

Levy, David. *Clyde Tombaugh, Discoverer of Planet Pluto*. New York: Sky, 2007. Written by an amateur comet hunter (of Shoemaker-Levy fame), this book details the detective story that was Clyde Tombaugh's quest at the Lowell Observatory in Flagstaff, Arizona, to find Planet X, which we now call Pluto.

Littmann, Mark. *Planets Beyond: Discovering the Outer Solar System*. New York: Dover, 2004. A history of exploration of the outer solar system as well as of Earth-based observations of Uranus, Neptune, and Pluto. Essays are interjected into the text profiling astronomers who have made important contributions.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. An introduction to the properties of all the major objects of the solar system, at the level of a college distribution course in astronomy. A good source of data about the solar system, taking a comparative planetology approach.

Tombaugh, Clyde W., and Patrick Moore. *Out of the Darkness: The Planet Pluto*. New York: New American Library, 1981. A classic about the discovery of Pluto by its discoverer, a well-known popularizer of astronomy. Tombaugh describes his career as well, and coauthor Moore describes the discovery of outer planets Neptune and Uranus. Of interest primarily for its historical perspective, since much of the material on the physical properties of Pluto and Charon is outdated.

Tyson, Neil deGrasse. *The Pluto Files*. New York: W. W. Norton, 2009. When the International Astronomical Union devised a new scheme for classifying solar-system objects, Pluto was suddenly demoted to dwarf-planet status, much to the displeasure of many astronomy buffs and large segments of the astronomical community. This book provides an account of Pluto's various classifications, from ninth planet to its loss of planetary status.

See also: Dwarf Planets; Eris and Dysnomia; Kuiper Belt; Nemesis and Planet X; Planetary Atmospheres; Planetary Classifications; Planetary Interiors; Planetary Orbits; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Solar System: Origins.

Protostars

Category: The Stellar Context

Protostars are stars in the process of forming. Planets are a natural by-product of the formation of stars, so protostars are most likely to be forming both stars and attendant planetary systems simultaneously. Understanding how other stars and planets form provides insight into how our own solar system formed.

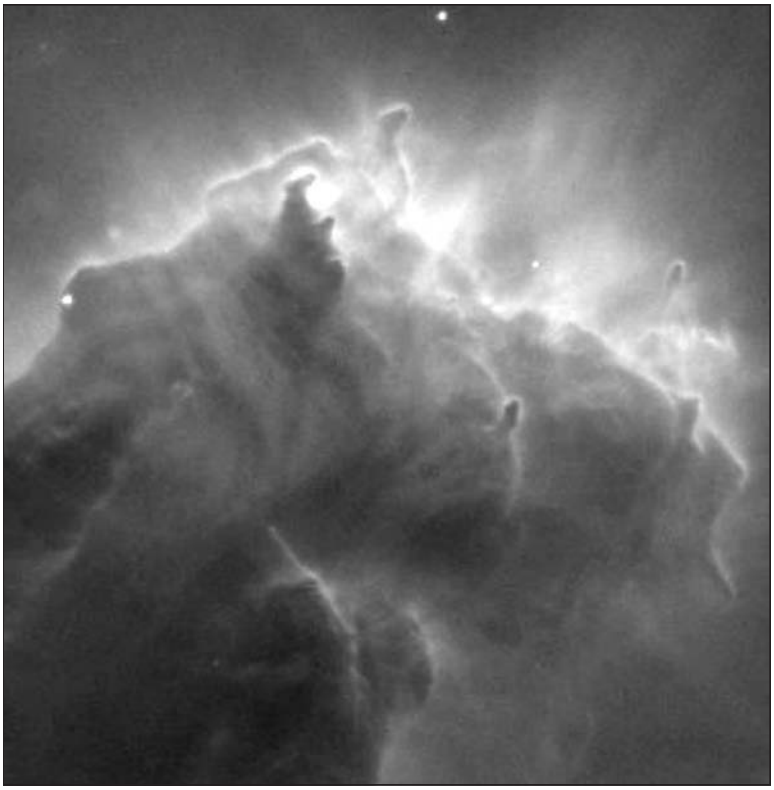
OVERVIEW

In the Northern Hemisphere, the constellation Orion dominates the early evening sky. If one scans the familiar sword and belt with binoculars or a small telescope, the center star in the sword looks like a small, fuzzy patch that resembles a piece of lint on the lens. This patch is actually the famous Orion nebula and is the closest and therefore best-studied star-forming region in the sky. More powerful telescopes reveal many similar nebulae where star formation is currently taking place.

Other stars and our own Sun formed by similar processes. A nebula, which is a cloud of gas and dust in interstellar space, begins to collapse. Once something triggers the collapse, the the nebula's own gravity takes over to continue the collapse. As the cloud collapses, it begins to spin more rapidly, like a figure skater bringing in her arms. The rapid spinning causes the cloud to flatten into a disk, like pizza dough when it is tossed spinning into the air. Most of the gas and dust in the nebula collapses into the protostar, which will become the central star. The leftover gas and dust in the surrounding disk will coalesce into planets orbiting the star. This process, known as the nebular hypothesis of solar-system formation, is the most widely accepted explanation of the origins of our solar system. In this way planets naturally form out of the debris surrounding the protostar. Therefore astronomers expect that most stars should be surrounded by an entourage of planets.

Astronomers can begin to understand the details of this process by studying the Orion nebula and similar nebulae. Beginning in the 1950's and 1960's, new wavelength regimes, notably radio and infrared, opened up to astronomers. Astronomers were now able to probe the dusty depths of star-forming nebulae by analyzing their radio and infrared emissions. Formerly these nebulae were "invisible" to astronomers becuase the dust found in these clouds blocks visible light; however, the nebular dust tends to allow more infrared and radio waves to get through, allowing the emissions to be collected by instruments sensitive to these regions of the electromagnetic spectrum.

When astronomers find a new tool to probe star-forming regions, they usually try using it first on the Orion nebula and then on other star-forming regions. At infrared wavelengths astronomers found that the Orion nebula contains several sources of infrared light. Similar infra-



This cloud of hydrogen gas in the Eagle nebula forms an incubator for new stars that are forming in its "fingertips." Each tip is larger than our own solar system. The pillars of gas will photoevaporate, revealing the evaporative gaseous globules (EGGs) that eventually may form stars. (NASA/ESA/STScl/J. Hester & P. Scowen, Arizona State University)

red sources are very common in other nebulae associated with star formation. The infrared signature of these sources tells astronomers that the sources are warm but not hot. The dusty cocoon out of which the star is forming still surrounds the protostar. The dust blocks the direct light from the protostar, but the energy from the protostar heats the dusty shell enough that it glows in the infrared. Hence infrared sources in star-forming nebulae are signatures of a protostar that is still enshrouded in a dusty cocoon.

At radio wavelengths, astronomers found both giant molecular clouds and H II regions associated with star-forming nebulae. The symbol H II is from a notation used by astronomers to indicate that the hydrogen atoms have lost their electrons and are therefore ionized. H II regions are therefore regions of ionized rather than neutral hydrogen. H II regions can surround newly formed stars, if the stars are hot enough to emit significant amounts of ultraviolet light. The stars falling into this category are massive O and B spectral class stars. As the H II region expands around these massive hot stars, the leftover material from the protostar stage is swept away. Hence H II regions allow massive protostars to get rid of their leftover cocoons of gas and dust.

Molecular clouds are also often associated with protostars. Typically they contain a few hundred thousand protostars, but giant molecular clouds can contain as much as ten million times the mass of the Sun and extend for hundreds of light-years. They are clouds of molecules in interstellar space. Astronomers think that the molecules form on the surfaces of interstellar dust grains and then escape the dust grains. More than a hundred different types of molecules have been identified in molecular clouds. The first stage in star formation is often a clump starting to collapse within a giant molecular cloud. These clumps will typically form into a small group of a half a dozen to a dozen protostars. As these protostars collapse, they heat up enough to ignite nuclear reactions in their cores. They then become stars and warm the surrounding dusty cocoon so astronomers see infrared sources. If the stars are massive stars, they emit enough ultraviolet light to form H II regions. In addition to clearing away the leftover material, the expanding H II regions send a shock wave into the giant molecular cloud and trigger the collapse of a new batch of protostars.

Massive stars produce H II regions to clear away the protostellar cocoons of leftover material, but how can less massive stars shed their cocoons? One way is the T-Tauri phase. T-Tauri stars have very strong stellar winds, which are similar to the solar wind but much stronger. Younger T-Tauri stars still have shells of surrounding material, while older (often called "naked") T-Tauri stars do not. If a less massive protostar goes through a T-Tauri stage, the strong stellar winds can blow away its shell of leftover material.

The Hubble Space Telescope has helped astronomers understand other ways less massive protostars can rid themselves of their shells of leftover material. Close-up photographs of the Orion nebula made with the Hubble telescope reveal protoplanetary disks (or proplyds for short). These proplyds look like small, dusty disks. The protostars are collapsing and spinning quickly, so they flatten into a disk. The central regions of these proplyds will become the stars, and the disks will coalesce into the planets. Nearby newly formed O and B stars can ionize the remaining hydrogen in these proplyds.

In 1995, the Hubble Space Telescope took a high-resolution photograph of the Eagle nebula, also known as M16. This strikingly beautiful photograph shows a large number of what astronomers call evaporative gaseous globules (EGGs). These EGGs are most likely the same thing as proplyds but seen from a different angle. In M16 there is a cluster of young hot stars just above the EGGs. Ultraviolet light from these stars helps rid the low-mass protostars forming in the EGGs of their excess material in a manner similar to the way massive protostars do with their own ultraviolet light. Other Hubble Space Telescope pictures of star-forming nebulae show similar phenomena, so it is believed that nearby massive stars provide the ultraviolet light that less massive stars need to shed their protostellar cocoons.

KNOWLEDGE GAINED

The earliest studies of star-forming regions such as the Orion nebula and similar nebulae were at visible wavelengths only. Protostars are still enshrouded in the dust out of which they formed, and this dust blocks visible light. Hence it is difficult to learn much about protostars by studying them only with visible light. A probe that penetrates dust is needed. The advent of radio and infrared astronomy provided the needed probe. Infrared light and radio waves have wavelengths that are longer than typical interstellar dust grains. Hence the dust grains do not block them as they do visible light. In the late 1960's, when these tools became readily available to the astronomical community, the study of protostars and star formation blossomed.

Most astronomers thought that molecules could not form in interstellar space, but the discovery of giant molecular clouds with radio telescopes proved them wrong. Because giant molecular clouds are the site of the initial stages of star formation, their discovery helped astronomers understand protostars. The study of radio H II regions also helped astronomers understand how very hot massive stars rid themselves of the gas and dust shells left over from their formation.

Infrared astronomy also gave astronomers important clues for understanding protostars. Just as red stars are cooler than blue stars, objects that are brightest at infrared wavelengths are cooler than stars. Dust shells surrounding protostars are warm but not as hot as stars, so infrared astronomy helped astronomers find stars and protostars in these early stages.

Launch of the Hubble Space Telescope gave astronomers the ability to take pictures of unprecedented resolution. High-resolution photographs of the Orion nebula, M16, and other star-forming regions gave astronomers direct views of proplyds and EGGs, which are solar systems in the process of forming. A next generation space telescope, when launched, will likely produce similar dramatic advances.

Astronomers have long thought that the planets in our solar system formed from the material left over from the Sun's formation and that planets are a natural by-product of star formation. Discovery of proplyds and EGGs support this theory, and if this theory is correct planets around other stars should be very common. In 1995 the discovery of the first extrasolar planet was announced. Since that time about two hundred extrasolar planets have been detected. Discovery of these extrasolar planets helps confirm astronomers' ideas that the solar system and other stars formed by the same process.

CONTEXT

One of the questions that virtually all young children ask is, "Where did I come from?" Parents usually interpret this question in terms of reproductive biology. However at a much deeper level this question concerns origins: the origin of life, the origin of Earth, the origin of the solar system, and the origin of the universe. Studying protostars helps us understand the origin of stars and, because the Sun is a star, the origin of the solar system that includes Earth. Hence studying protostars and star formation relates to the universal human quest to understand where we came from. The study of origins has been a strong focus for the National Aeronautics and Space Administration (NASA).

Because protostars are one of the earliest stages in the life cycle of stars, studying protostars is part of the larger study of stellar evolution. To completely understand the entire life cycle of stars, it is necessary to understand each of the stages a star goes through, including the protostar stage. Only the lightest elements on the periodic table were made during the big bang. The carbon, oxygen, and other atoms important to life were made later in stars and blasted back into space to be recycled by the next generation of protostars. Hence, the study of stellar evolution in general relates to the question of human origins, the origin of the atoms in our bodies.

Paul A. Heckert

FURTHER READING

Bally, John, and Bo Reipurth. *The Birth of Stars and Planets*. Cambridge, England: Cambridge University Press, 2006. This lavishly illustrated book is written by two astronomers who specialize in star formation.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. This well-written introductory astronomy textbook completely discusses star formation as well as stellar evolution and the formation of the solar system.

Cohen, Martin. *In Darkness Born: The Story of Star Formation*. Cambridge, England: Cambridge University Press, 1988. Written by a researcher in star formation, this is a readable account of the story of star formation, though it was published before the results of the considerable data resulting from the Hubble Space Telescope.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. The third major section of this well-written introductory astronomy textbook contains seven chapters on stars and stellar evolution, including a chapter on star formation. The section on the solar system also contains a chapter on the origin of the solar system.

Hester, Jeff, et al. *Twenty-First Century Astronomy*. New York: W. W. Norton, 2007. Chapter 14 of this readable astronomy textbook is about star formation; subsequent chapters cover the rest of stellar evolution.

Inglis, Mike. *Observer's Guide to Stellar Evolution*. London: Springer, 2003. Chapter 2 of this book is on star formation. The other chapters tell the entire story of the lives of stars. An unusual feature of this book is a guide that amateur astronomers can use to observe the various stages of stellar evolution.

Lin, Douglas N. C. "The Genesis of Planets." *Scientific American* 298, no. 5 (May, 2008): 50. This article describes the latest knowledge about the formation of planets around protostars.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. Cambridge, England: Cambridge University Press, 2002. An introductory astronomy textbook that is extremely well written, taking into account research on how students learn astronomy. There is a chapter on star formation as well as chapters on the formation of the solar system and stellar evolution.

See also: Hertzsprung-Russell Diagram; Infrared Astronomy; Interstellar Clouds and the Interstellar Medium; Main Sequence Stars; Planetary Formation; Red Giant Stars; Stellar Evolution.

Pulsars

Category: The Stellar Context

Pulsars were unexpectedly discovered in 1967. These objects emit electromagnetic radiation that pulses on and off with periods ranging from approximately 0.001 second (one millisecond) to 4 seconds. Pulsars are thought to be rapidly rotating neutron stars, the remnants of massive stars which gravitationally collapse to a size of only about 10 kilometers after their progenitor star first goes supernova.

OVERVIEW

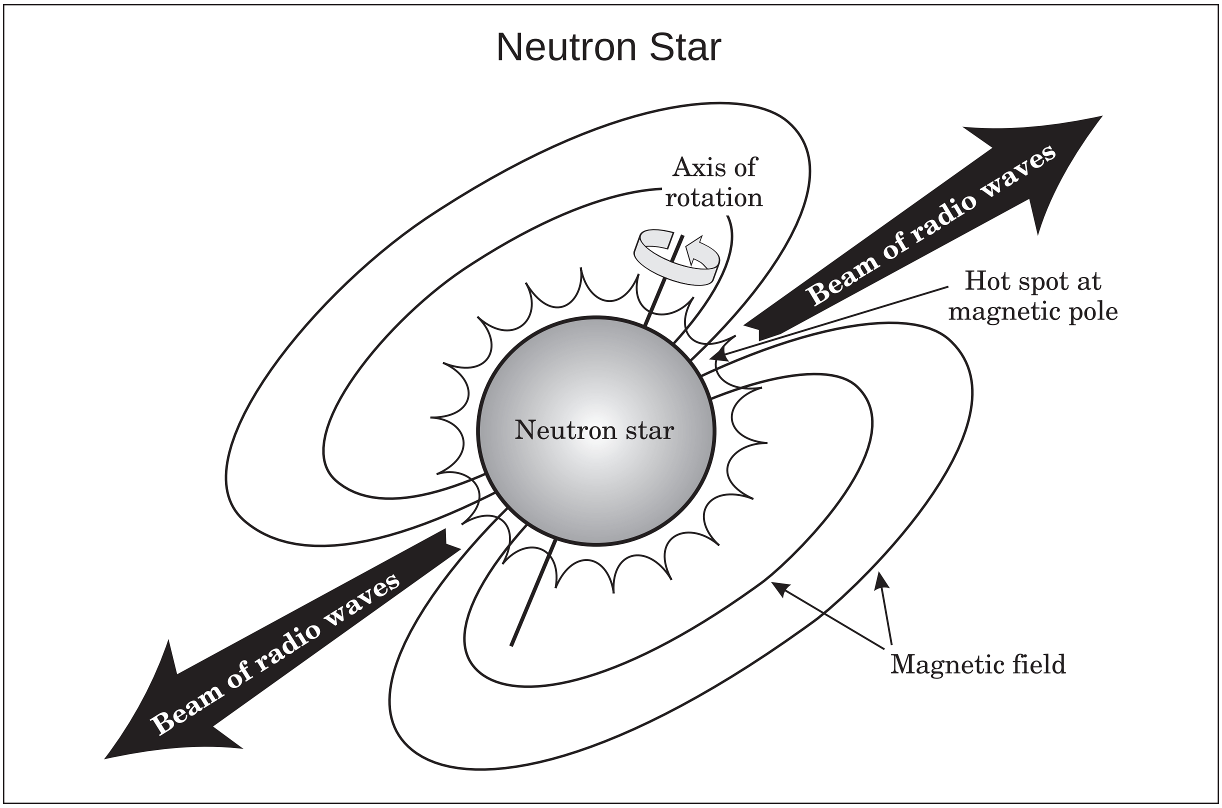
Pulsars were accidentally discovered in 1967 by Jocelyn Bell (later Jocelyn Bell-Burnell), while working with a group of British radio astronomers led by Antony Hewish. Other members of the group were John D. H. Pilkington, Paul Frederick Scott, and Robin Ashley Collins. The Hewish group was looking for interplanetary scintillation at radio frequencies. Scintillation is the twinkling of radio sources caused by interplanetary ionized gas.

On August 6, 1967, they noticed a mysterious signal that looked like random interference, such as that from a passing automobile. The signal reappeared at the same location in the sky, however, and therefore had to be of celestial origin. Using a high-speed recorder, researchers then observed this signal (from what was later named CP 1919) on November 28, 1967. The high-speed recorder showed extremely regular pulses (brief intervals of greater than normal brightness). Timing was repeatable to one part in about one million. That extreme regularity caused the Hewish group to consider briefly the possibility that they were observing a evidence of extraterrestrial intelligence. This conjecture was dubbed the “Little Green Men” (LGM) theory. Subsequent discovery of three additional similar signals forced the Hewish group to consider natural origins for their mysterious signal since similar multiple sources would not likely all be from intelligent extraterrestrial civilizations.

The Hewish group named this newly discovered class of astronomical objects pulsars, because they seemed to be stars pulsing on and off very rapidly. The first pulsar discovered by the Hewish group was designated as CP 1919. CP 1919 and other pulsars are named by their position in the sky. In early pulsar nomenclature, CP stands for “Cambridge Pulsar,” the University of Cambridge being the location of the discovery group. The 1919 refers to the object’s right ascension, which is one of the position coordinates in the sky and analogous to longitude. In modern nomenclature, pulsars are designated by PSR, short for “pulsar,” and a number indicating the pulsar’s location in the sky. Particularly famous or important pulsars are also given names. For example, PSR 0531 is also known as the Crab nebula pulsar because it is centrally located within the Crab nebula.

Like many discoveries in science, the discovery of pulsars was serendipitous. The Hewish group was not looking for pulsars and indeed did not even suspect that pulsars existed. Their equipment, however, was designed to measure the angular structure of radio sources using interplanetary scintillation. They were thus able to detect rapid variations of pulsars. As with all unexpected discoveries, there was a temptation to disregard data not conforming to prior expectations and suspect such data to be the result of either interference or an idiosyncrasy of the equipment. Other groups had in fact observed pulsars at both radio and optical wavelengths prior to the Hewish group’s discovery, but they did not recognize the significance of their observations until the Hewish group announced their findings. Credit for following up on the unexpected data, searching for their extraterrestrial origin, and thereby discovering pulsars justly goes to Bell-Burnell.

Pulsars were originally discovered at radio wavelengths, but they have since been observed over the entire electromagnetic spectrum. W. John Cocke, Michael J. Disney, and Donald J. Taylor, working at the University of Arizona, first discovered a pulsar pulsing at optical wavelengths on January 16, 1969. This pulsar, the famous Crab nebula pulsar (PSR 0531), has also since been found to pulse over the entire electromagnetic spectrum, from X-radiation to radio waves. Pulses from the Vela pulsar have also been recorded over a wide wavelength



A diagram of a neutron star showing its strong magnetic field, which generates radiation that can be detected on Earth as radio waves.

range. Pulsars seem to pulse at all wavelengths; this fact has been directly observed, however, for only a small fraction of pulsars. This discovery is significant, because pulsars must produce much more energy to pulse at optical or X-ray wavelengths than at radio wavelengths only.

Since the initial discovery, astronomers have continuously searched for new pulsars and have extensively studied the properties of known pulsars. By 2008, the pulsar catalog listed more than sixteen hundred identified objects; the total number of pulsars had doubled during the previous decade. Estimates based on current astrophysical understanding suggest that as many as half a million pulsars should exist in the Milky Way, but most are not readily detectable. No pulsars have yet been detected outside our galaxy, but pulsars cannot be objects confined to existence only in the Milky Way of course. The capability to detect pulsars is limited by their relatively low luminosities.

Pulsar periods, the time required for a complete on-off-on pulse cycle, range from about a millisecond (0.001 second) to almost 4 seconds. These periods are very regular and precise, to about one part in 100 million. In addition, periods for the faster pulsars, notably the Crab nebula pulsar, are known to be increasing—that is, their pulse rates are slowing down. Pulsars also have occasional "glitches," or abrupt changes in their periods, which are thought to be caused by "starquakes" on the pulsar.

Each pulsar has its own unique pulse shape. Some pulsars have two distinct pulses, a major pulse and an interpulse. The interpulse is usually smaller than the main pulse, but can be as large. Other pulsars have only one pulse. An individual pulsar's pulse shape over a single cycle can vary, but the average pulse shape over a large number of cycles remains constant. No two pulsars are known to have the same average pulse shape.

Pulsars are also known to be associated with supernova remnants. A supernova remnant is formed from the material blown away from a massive stellar explosion. For example, the Crab nebula is the remnant of a supernova that was observed in 1054. The Crab nebula pulsar is associated with this remnant and is thought to be the collapsed core of the star that went supernova.

A theoretical model explaining pulsars must account for all observed properties. The Hewish group originally thought that pulsars changed in brightness by increasing and decreasing in size, or pulsating. Hence, they were named pulsars. However, this pulsation was not able to account for all the later-observed properties of these objects. The Hewish group did correctly conclude that pulsars were most likely to be either white dwarfs or neutron stars.

Thomas Gold suggested in 1968 that pulsars are actually rotating neutron stars rather than pulsating stars. According to his "lighthouse model," the pulsar sends out relatively tight beams of radiation in two directions. As the pulsar rotates, the beam alternately points toward and away from Earth; thus the pulsar appears to flick on and off rapidly. An individual pulsar's pulse shape is determined by how squarely these beams hit Earth. If both beams hit Earth equally, two equal pulses are recorded. If one beam misses, only one pulse is detected. If one beam hits Earth dead center and the other barely hits, a strong main pulse and a smaller interpulse result. If both beams miss, the pulsar cannot be detected from Earth. Thus, it is likely that there are pulsars that cannot be identified from Earth. Whether a pulsar and its pulse shape are detectable depends on the chance alignment of these beams relative to Earth.

Pulsars are rotating neutron stars rather than white dwarfs because pulsars must be very compact, and the rotation periods, as given by the pulse times, are too fast for objects as large as white dwarfs. A typical white dwarf star will have about the radius of Earth, which is too large to explain the rotation periods of the faster pulsars. A neutron star, on the other hand, with a typical radius of about 10 kilometers, depending on its mass, is small enough for even the most rapid pulsar.

Why is the pulsar/neutron star spinning so fast? Neutron stars are formed when a massive star explodes in a supernova. After the supernova, the core of the star collapses to form a neutron star. If the original star was rotating, the collapsed core turned neutron star will continue to rotate much faster—many times a second. For an illustration of the mechanics of this phenomenon, consider the figure skater who wants to spin: She must pull her arms close to her body in order to spin faster. To slow down, she stretches out her arms. Similarly, as the stellar core collapses into a 10-kilometer neutron star, it begins spinning very rapidly. The spinning property known as angular momentum will be conserved, or remain constant, if there is no net external torque. Because the amount of angular momentum depends on the rotational radius as well as the rotation rate (the mass rotating is also important), as the radius decreases the rotational rate increases. The neutron star must spin very rapidly because it is so small compared to the original star.

As the neutron star collapses, it also compresses the star's magnetic field. Hence, there is a very strong magnetic field at the surface. Electrons moving at near the speed of light in a strong magnetic field cause what is known as synchrotron radiation to be emitted. The exact mechanism is still poorly understood, but the magnetic field of the pulsar causes two beams of synchrotron radiation to be emitted from the pulsar's magnetic poles. If the pulsar's spin and magnetic axes are lined up in such a way that one or both of the beams are pointed toward Earth, a pulsar is observable. The orientation of these two axes will determine the exact pulse shape that is observed. Ultimately, the energy observed from pulsars and the surrounding supernova remnant comes from the rotation of the pulsar. If the model is correct, then pulsars must slow down. The observed increase in pulsar periods strongly supports the lighthouse model.

The model of pulsars as rapidly rotating neutron stars, then, explains the observed properties of pulsars. The fact that neutron stars are thought to be formed in a supernova explosion and that several pulsars are associated with the remnants of a supernova explosion strengthens this connection between pulsars and neutron stars.

KNOWLEDGE GAINED

During the 1930's the possibility of neutron stars was first hypothesized by physicist J. Robert Oppenheimer, but prior to 1967 they had not been observed. With the discovery and later explanation of pulsars, neutron stars were shown to exist. Their discovery provided an important corroboration of accepted theories of stellar evolution.

However, neutron stars did not have exactly the properties that had been predicted. Few if any astronomers had suspected that neutron stars would wink on and off in less than a second. Franco Pacini perhaps had come the closest to predicting, prior to the discovery of pulsars, this true nature of neutron stars. In 1967, Pacini correctly predicted that neutron stars should be rotating rapidly and suggested that this rotation in a condensed magnetic field provided the energy for the surrounding supernova remnant. In trying to explain the then unknown energy source within the Crab nebula and other supernova remnants, Pacini came remarkably close to predicting the existence of pulsars, but he did not predict the beams of radiation that constitute pulses from the pulsar.

The lighthouse model, explaining pulsars as rotating neutron stars, seems obvious now but had not been considered prior to the discovery of pulsars. Even after their discovery, most researchers thought that pulsars were white dwarfs rather than neutron stars, until pulsars too fast to be white dwarfs were discovered. Originally, most astronomers thought that pulsars were more likely to be white dwarfs because the latter had been observed for quite some time. For an object the proposed size of neutron stars to be able to rotate as fast as the observed pulsars yet not fly apart, the density of that object must be much greater than a value typical of the density of white dwarf material. This physical constraint, coupled with the discovery of the faster pulsars, convinced astronomers that pulsars must be neutron stars rather than white dwarfs.

After observing the properties of pulsars, most astronomers agreed that pulsars were indeed neutron stars and realized that pulsars provided an excellent laboratory for the study of a number of phenomena. Neutron stars are a form of highly condensed matter: degenerate neutrons. Degenerate neutrons are compressed to roughly the density of the atomic nucleus, so that if the neutrons were compressed any further they would lose their structure as neutrons. These conditions cannot be duplicated in a laboratory on Earth, so to understand how matter behaves under these conditions scientists must study pulsars. A pulsar shines by synchrotron radiation, which is produced by the interaction of high-speed electrons moving in an intense magnetic field. The scale and intensity of a pulsar's magnetic field cannot be reproduced on Earth, so by studying pulsars researchers can gain information about regions with strong magnetic fields.

A pulsar also provides the energy source that illuminates the supernova remnant surrounding the pulsar. An example is the Crab nebula. Prior to the discovery of the Crab nebula pulsar, astronomers did not understand the energy source for this nebula. It is now known that the pulsar provides the energy to make the nebula shine. The same mechanism can cause other supernova remnants to shine similarly.

CONTEXT

Prior to the discovery of pulsars, scientists had suggested that neutron stars might be formed from the stellar core left after a supernova. Once pulsars had been discovered to be associated with supernova remnants, this suggestion had a firm observational foundation. Since the discovery of pulsars, astronomers better understand the formation of neutron stars and the role of supernovae in it. The first supernova observable to the naked eye in nearly four hundred years was observed in the late 1980's in the Large Magellanic Cloud (Supernova 1987A). Scientists were hoping that a pulsar would eventually form, affording them an opportunity to observe directly the formation of a neutron star.

Placed in a broad context, the most significant aspect of the study of pulsars is their role in increasing astronomers' understanding of stellar evolution. In particular, pulsars have helped clarify the role of supernovae in the process of star death. It is thought that neutron stars and black holes are the final corpses of massive stars and are formed in supernova explosions. The study of pulsars increases scientists' understanding of supernovae and ultimately of Earth's origin.

Supernovae play a crucial role in the manufacture of the raw materials needed for life. According to the big bang model, in the aftermath of the universe's formation it contained primarily hydrogen and helium with trace quantities of lithium and beryllium and no other elements. How, then, are all the other naturally occurring elements formed? Stars manufacture other elements by nuclear fusion reactions that produce stellar energy. For example, stars similar to the Sun can produce elements as heavy as carbon in their cores before they collapse into white dwarfs. Stars that are much more massive than the Sun can produce elements as heavy as iron in their cores, and elements slightly heavier than iron nearer the surface. The heaviest elements probably cannot, however, be manufactured in stellar cores, because the processes by which elements heavier than iron are manufactured require rather than release energy. Supernovae release tremendous quantities of energy, so they can supply all the energy needed to manufacture the heavier elements.

Supernovae also provide the recycling mechanism for the elements manufactured in stellar cores. Such an explosion blows material rich in heavy elements back into space, where it can be recycled in the next round of star and planet formation. Without supernovae, this material would be trapped in stellar cores. With the exception of hydrogen and helium, all the material on Earth and in human bodies was manufactured in the core of a star or in a supernova and blown into space by a supernova, prior to the formation of the solar system.

By studying pulsars, scientists gain a greater understanding of the supernova process and the role played by supernovae in manufacturing the raw materials necessary for human existence. Without clues provided by pulsars, scientists could not have unraveled the mystery of stellar death and its essential role in the development of life on Earth.

Paul A. Heckert

FURTHER READING

Condon, J. J., and S. M. Ransom. *Essential Radio Astronomy*. Available online at http://www.cv.nrao.edu/course/astro534. Detailed notes from a National Radio Astronomy Observatory/University of Virginia course in radio astronomy. Includes links to chapters on pulsar physics and observations.

Gurevich, A. V., V. S. Beskin, and Ya N. Istomin. *Physics of the Pulsar Magnetosphere*. Cambridge, England: Cambridge University Press, 1993. A thorough review of our contemporary understanding of the electrodynamic phenomena associated with the magnetosphere of pulsars. Technical.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find the introductory texts pitched at too low a level. Suitable for an audience with varied science and mathematical backgrounds. Covers all topics from solar-system objects to cosmology.

Lyne, Andrew G., and Francis Graham-Smith. *Pulsar Astronomy*. 3d ed. Cambridge, England: Cambridge University Press, 2006. An introduction to the physics of pulsars, their unique electromagnetic emissions from X rays to radio waves. Technical.

Ostriker, Jeremiah P. "The Nature of Pulsars." *Scientific American* 224 (January, 1971): 12, 48-60. Written shortly after pulsars were fairly well understood, this classic summarizes the observed characteristics of pulsars.

Pulsating Stars. 2 vols. Introductions by F. G. Smith, A. Hewish, and T. Gold. New York: Plenum Press, 1968-1969. A two-volume set containing reprints of early articles from *Nature* on pulsars. Many of the papers on the discovery of pulsars were printed in *Nature*, so these volumes make a good collection on the historical development of pulsar studies.

Sion, Edward M., Stephane Vennes, and Harry L. Shipman. *White Dwarfs: Cosmological and Galactic Probes*. New York: Springer, 2005. Suitable for college students. Discusses the cosmological importance of white dwarfs and the physics of their formation and ultimate fate.

Smith, F. G. *Pulsars*. Cambridge, England: Cambridge University Press, 1977. A good review of knowledge of pulsars through the late 1970's. The first chapter provides a nontechnical review of the discovery of pulsars. Later chapters provide a moderately technical summary of observations and understanding of pulsars.

Verschuur, Gerrit L. *The Invisible Universe: The Story of Radio Astronomy*. New York: Springer Praxis, 2006. Provides a history of developments in radio astronomy, and along the way describes the discovery of pulsars, quasars, and radio galaxies. Suitable for a general science course in college as well as for astronomy majors as background information.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. New York: John Wiley and Sons, 2002. An excellent introductory astronomy text. The section on pulsars is comprehensive and well written. The rest of the section on stellar evolution places the topic of pulsars in a broader context, while the entire text places stellar evolution in context. The treatment assumes no prior knowledge of astronomy, so this volume is suitable for a general audience.

See also: Extrasolar Planets; Extrasolar Planets: Detection Methods; Gravity Measurement; Novae, Bursters, and X-Ray Sources; Planetary Formation; Radio Astronomy; Stellar Evolution; Supernovae.

R

Radio Astronomy

Category: Scientific Methods

In 1932, Karl Guthe Jansky made observations that led him to conclude that Earth was receiving radio waves not produced by humans. This event opened a new observational area of study for astronomers and gave birth to the science of radio astronomy, which produced a new picture of the universe as a very active place with pulsars, quasars, and radio galaxies.

OVERVIEW

Until radio experiments by Karl Guthe Jansky (1905-1950) in the 1930's, virtually all astronomical knowledge was a product of analyzing visible starlight. Historically, optical telescopes and photographic techniques enhanced scientists' vision of the small range of frequencies in the electromagnetic spectrum perceived as visible light (400 to 700 nanometers in wavelength). Radio signals are found in the long-wavelength part of the electromagnetic spectrum, ranging from millimeters to several thousand meters. Jansky's 1932 discovery of radio sources beyond Earth ultimately led astronomers to imagine millions of discrete radio transmitters broadcasting throughout the universe. The small slit of visible light through which astronomers had previously viewed the cosmos suddenly expanded to a wide-open window with the addition of radio signals. Unfortunately, the radio universe is invisible to humans' biological eyes, and consequently it is necessary to convert radio signals into a visual format for interpretation of data.

A radio telescope collects and converts radio signals through its basic components: the antenna, which collects the signals; an amplifier, which magnifies these incredibly weak waves; and a recorder, which translates this information into a medium that can be viewed. A radio telescope is a directional antenna that may take several forms. Jansky's design was a square wooden frame around which he wrapped wire. Another type of antenna is the dipole, a pair of metal poles (the length of the wavelength to be received) 180° to each other joined by insulating material. A television antenna is a common example. Television antennas often have several dipoles in front of the collector element (where the wires attach from the television) that focus its directivity, and a reflector element behind it to increase the strength of the incoming signal. Radio telescopes can be very similar, with only the length of their dipoles being different. The size, shape, and design of a radio telescope is in large part dictated by the radio-wave frequency being investigated by the telescope.

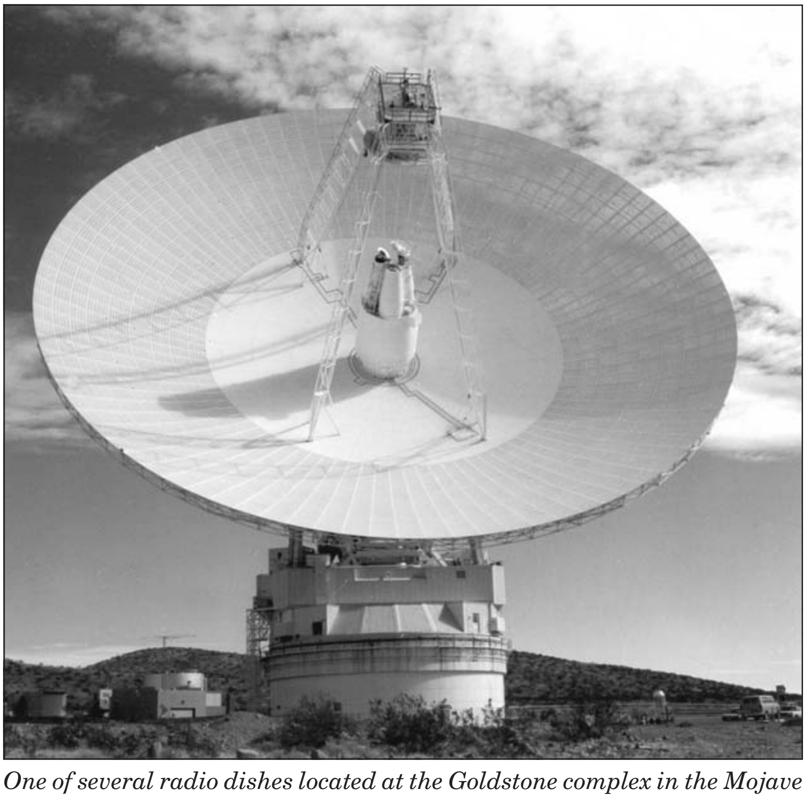
Nevertheless, as signals from the extraterrestrial sources are many times weaker than television broadcasts, radio astronomers must use antennae of the greatest directivity and sensitivity. The parabolic reflector efficiently satisfies these requirements. For radio astronomy, these reflectors are generally larger versions of the home television satellite dish.

While the advantage of a larger observational window is attractive, radio astronomy has its limitations. The ability to see, or resolve detail, with a radio telescope depends on the size of its reflecting elements, proportional to its diameter, very much like the resolution parameters of optical telescopes. Radio waves, longer than optical wavelengths by ten thousand, require reflectors to be equally large. Since this requirement is impractical to achieve, most radio telescopes have, by optical standards, "poor" resolution. A parabolic reflector 30 meters in diameter has a resolving power of about 0.5°, or about the size of the full Moon. Fortunately, advances in computers allow radio telescopes kilometers apart to be linked together, forming the equivalent of a collector equal to that distance.

Another difficulty facing radio astronomers is the extremely weak nature of the signals. Even large telescopes such as the 300-meter dish at Arecibo, Puerto Rico, collect only minute quantities of radio signals. To illustrate the incredibly faint nature of these signals, one must realize that the energy received by all radio telescopes on Earth is equivalent to less than the energy transferred by a snowflake hitting the ground. Additionally, radio astronomers have to filter unwanted radio signals created by their own technology. Human-produced radio noise can range from automobile ignitions to satellite transmissions.

As severe as these limitations may appear, radio astronomy provides some distinct advantages over optical telescopes. Radio telescopes are not affected by weather, light pollution, interstellar dust, or the Sun, which obliterates observations for optical telescopes. The ability of radio telescopes to identify optically invisible sources of matter in between the stars, such as nonluminous clouds of hydrogen, makes them an invaluable tool. Hydrogen emits a distinctive radio signature originating from a spin-flip transition in the ground state that emits photons of 21-centimeter wavelength (or a frequency of 1,420 megahertz). The universe is dominated by hydrogen (about 75 percent), and the study of its distribution contributes to an understanding of the cosmos. Additional advantages of the radio window are the property of these long wavelengths to penetrate optically opaque dust clouds in the Milky Way and allow astronomers to "see" into the Galaxy. The radio window allows astronomers to detect objects that are "brighter" in the radio frequencies than in the visible wavelengths. Indeed, some of the most distant objects in the universe are detectable only by their radio emissions.

Radio waves are generated in space by several methods. Thermal emissions—frequencies radiating from hot ionized gas clouds—indicate the places of star birth. However, most radio emissions are of nonthermal origin. Cosmic-ray electrons (charged particles) ejected from supernova explosions are typical of this type of radio emission. Upon encountering magnetic fields, such as a galaxy, these electrons move in helical paths around and within the magnetic field, losing kinetic energy that radiates into space as radio waves. This source of radiation yields information about cosmic rays and the nature of interstellar magnetic fields. Radio waves also originate by the ground state spin-flip transition of hydrogen atoms scattered throughout interstellar space. The frequency is so specific that it has become a subdiscipline in radio astronomy: the study of the 21-centimeter emission line, hydro-



One of several radio dishes located at the Goldstone complex in the Mojave Desert of California. These sensitive paraboloid dishes can be moved to face incoming radio waves, which are reflected to a subreflector located in the housing at the top of the four scaffold-like towers (center), which in turn beams these concentrated waves through a feed horn in the center of the dish and down to a control room where a receiver amplifies the signals and computers render them into images. (NASA) gen's radio signature. The importance of this radio signature is that, by utilizing the Doppler shift, it is possible to determine if a hydrogen nebula is moving toward or away from Earth. Hydrogen atoms also absorb radiation from more distant sources. By studying the combined effects of absorption and transmission, it can be determined if the radio source lies beyond the hydrogen cloud. In this way, studies in the 21-centimeter wavelength have helped determine much of the structure of the galaxy.

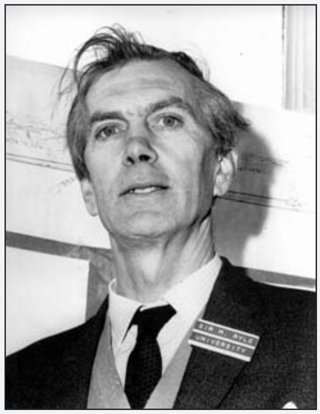
Earth's atmosphere is opaque to most radiation with the exception of the visible and radio wavelengths. The atmosphere absorbs most of the other wavelengths before they reach the ground. Radio telescopes enable astronomers to detect radio emissions from many sources: the Moon, Venus, Mars, Jupiter, Saturn, the Sun's atmosphere (corona), clouds of ionized hydrogen within the Milky Way, and cosmic-ray electrons spiraling in the magnetic field of the Milky Way as well as other galaxies. Astronomers have come to know much of the nature and structure of the Milky Way and the universe through radio observations.

APPLICATIONS

The universe, before the introduction of radio astronomy, appeared rather stable if not static apart from the occasional supernova. Then, with the development of radio and the associated radio telescope in the early twentieth century, a new picture emerged of an active universe. Discovery of radio sources identified as pulsars underlies this dynamic element.

Ryle's Radio Telescopes

In 1964, Sir Martin Ryle implemented his first history-making radio telescope, the "one-mile telescope," using the principle of aperture synthesis. Aperture synthesis uses small telescope dishes to produce the angular resolution of a much larger telescope dish. A telescope's angular resolution is its ability to distinguish between two relatively close point sources of radiation. The method of aperture synthesis keeps one or more small dishes fixed and moves one or more other dishes over Earth's surface, comparing the phases of the radiation collected by the fixed and movable dishes. Ryle's instrument was unique because he accounted for the effects of Earth's rotation in moving the array of dishes and provided a baseline for the angular resolution which was a large fraction of Earth's diameter.



(AP/Wide World Photos)

The telescope had a resolution superior to that of existing instruments, and it could detect much fainter sources, including quasars. Quasars are among the most distant, and therefore youngest and most powerful, objects in the universe. They may represent the stage galaxies go through before their radio radiations subside and they become visible at optical wavelengths. With Allan Sandage, Ryle developed a technique for identifying quasars at optical wavelengths, based on the fact that they emit much more ultraviolet energy than single, normal stars.

Ryle's survey of radio sources showed, as suspected, that the number of faint sources per unit volume of space increased with distance—but far more rapidly than anticipated. This finding supported the big bang theory of the universe, which then was in conflict with the prevailing "steady state" theory. Over time, Ryle's evidence was bolstered by other research, and today the big bang theory is dominant.

In 1974, Ryle won the Nobel Prize in Physics for construction and use of the five-kilometer telescope. Like the one-mile telescope, it consisted of a linear, east-west array of dishes, some fixed and some movable, all carried by Earth's rotation. The resolution of this instrument was, remarkably, one second of arc. The new telescope detected fainter and more distant galaxies with greater precision, allowing many more objects to be used for statistical studies and the radio and optical components of the most distant objects to be matched. The five-kilometer telescope was also used to study individual stars in the Milky Way that were just being born. Ryle's telescopes thus opened a new window on the universe, revealing previously undetectable objects and insights.

In 1934, two astronomers, Walter Baade and Fritz Zwicky, proposed that the remaining stellar core from a supernova explosion represents the transition of a star to a neutron star. This theory received little attention until 1967.

In the mid-1960's in Cambridge, England, Antony Hewish and his colleagues were building a radio telescope to observe scintillation, or twinkling, of radio sources known as quasars. By coincidence, the parameters for the telescope matched the characteristics of pulsars (later to be described as rapidly rotating neutron stars). Jocelyn Bell, a graduate student, detected a regular pulsating source of radio emissions that kept appearing from the same area in the sky and with a regularity of every 1.33 seconds. Bell's discovery of pulsars was the beginning of a small revolution in astronomy. Their signals were unlike any previously detected from stars or galaxies. The intriguing idea of the Little Green Men (LGM) phenomenon was entertained for a brief time. Then another pulsar was discovered, which was pulsating at 1.2 seconds. It was the regularity of the pulses that distinguished them from the other celestial radio sources. As astronomers rushed to map the radio sky, they found several more pulsars within the next year. There are currently more than 1,500 known sources. The fastest pulsars were associated with supernova remanents (nebulae); the slower ones had no visible nebula.

The evolution of a pulsar began in the fires of a supernova explosion. As the star's outer shell expanded into space (the nebula), the rapidly spinning core pulsed radio signals and slowed. There was a correlation between the rate of spin and the presence or absence of nebula. The twenty-thousand-year-old Vela pulsar has a rate of about one second, while the one-thousand-year-old Crab nebula has a rate of about thirty-three pulses per second. The Baade-Zwicky neutron star theory was revived and developed momentum as a model to explain the phenomena. Indeed, pulsars seem to be rotating neutron stars, the final products of supernova explosions.

Interestingly, no one has actually seen a neutron star; they are too small to be seen at the distances where they present themselves. Were it not for their radio pulses and the ability to detect them, astronomers would remain ignorant of the relationship of supernovae and neutron stars. However, using very long baseline interferometry (VLBI), astronomers have been able to identify differences across the surface of neutron stars.

Although the LGM theory did not last long after Bell's first pulsar discovery, it did begin a series of observations and experimental transmissions through which it is hoped to detect the existence of extraterrestrial intelligence, known as the Search for Extraterrestrial Intelligence (SETI). The search for signals from other civilizations began in 1959 at the Green Bank National Radio Observatory in West Virginia. In 1974, at the dedication of the radio telescope at Arecibo, Puerto Rico, astronomers transmitted the first message from Earth intended for other civilizations.

The discovery of pulsars and neutron stars had many consequences for astronomy and radio astronomy specifically. Fundamentally, it caused physicists to reassess the states of matter to accommodate a new phenomenon of "degenerate matter." Radio astronomy has enabled astronomers to "listen" to the background radiation of the big bang; determine the nature of quasars and their probable relation to black holes; discover molecular hydrocarbons in space, some of the fundamental building blocks of life; and begin an active search for intelligent extraterrestrial life. Radio astronomy has appreciably increased astronomers' knowledge of the Milky Way and the universe.

CONTEXT

The ability to see the universe through the radio window is a fairly recent technological development, whose evolution parallels the maturation of radio and its associated technology. Transmission of radio waves through space was demonstrated in 1887 by German physicist Heinrich Rudolph Hertz (1857-1894) in experiments based on predictions of James Clerk Maxwell (1831-1879). This successful demonstration began the age of radio.

Some astronomers suspected that stars might be a source of radio waves; the most easily examined star of course is the Sun. The search for solar radio signals began in 1890 with Amer-

A long view of several of the radio dishes at the Goldstone complex, one of the three complexes that form the Deep Space Network (DSN). This system uses very long baseline interferometry to resolve radio signals received around the world. Radio telescopes kilometers apart can be linked together, forming the equivalent of a giant collector equal to that distance and hence providing higher resolution. The other complexes in the DSN are located near Madrid, Spain, and Canberra, Australia. (NASA)

ican inventor Thomas Alva Edison (1847-1931), followed in England by Sir Oliver Joseph Lodge (1851-1940) in 1894 and Charles Nordman of France in 1900. Negative results from their experiments diminished interest in this line of research. The idea of extraterrestrial radio waves generally receded from the minds of the astronomical community at this point. Scientists of the period were not aware of the ionized reflecting layer in the upper atmosphere (ionosphere), which effectively filters radio frequencies longer than about 20 meters.

This apparent failure of radio science did not deter its technological development. Radio caught the imagination of the public and scientists with Guglielmo Marconi's (1875-1937) transatlantic transmission in 1901. This achievement led Arthur Edwin Kennelly (1861-1939) and Oliver Heaviside (1850-1925) in 1902 to propose an electrified layer in the upper atmosphere (ionsphere), which reflected radio waves around the curve of Earth. Research into this new technology and "atmospherics" intensified during World War I. By the 1920's, radio research was flourishing in the major universities and especially in commercial laboratories such as Bell Telephone Laboratories in the United States and Marconi Telegraph Company in England.

During exploration of radio technology in the 1920's, meteorological conditions, such as lightning, were known to affect the quality of radio transmissions, but details of atmospheric disturbances and the Kennelly-Heaviside layer (ionosphere) were not understood fully. Nevertheless, radio technology advanced, and in 1929, shortwave transatlantic radio telephone service became available to the public. Unfortunately, the technology of the 1920's also introduced its own source of radio noise—specifically, the automobile and the intrinsic noise of the vacuum-tube receivers. Research into atmospheric radio noise led by Sir Robert A. Watson-Watt (1892-1973, the inventor of radar) in the 1920's began a sequence of studies resulting in the foundation of radio astronomy.

Karl Guthe Jansky (1905-1950) joined Bell Telephone Laboratories in 1928. His assignment was to build a radio receiver to record and study the intensity of shortwave atmospheric interference ("atmospherics") of about 15 meters, a frequency used for ship-to-shore and transatlantic communication. Jansky observed three types of static in his receiver. He identified two types of static as atmospheric disturbances, but the third was a very steady type of hiss-static. This steady hiss was the most intriguing because the strongest part of the signal advanced in time throughout the year. At first, Jansky thought it might be connected with the Sun, but radio recordings of a partial solar eclipse of the Sun on August 31, 1932, indicated that the Sun was not the source. (This period was a time of sunspot minimum. Jansky might have recorded solar radio emissions had the Sun been in a sunspot maximum phase.)

Jansky spent 1932 analyzing the radio signals and trying to distinguish a single source and the apparent angle of arrival. He observed that the maximum signal strength coincided with the passage of the Milky Way in front of his antenna. On June 27, 1933, at the Institute of Radio Engineers' Eighth Annual Convention in Chicago, he presented a paper entitled "Electrical Disturbances Apparently of Extraterrestrial Origin." He described the source as having a periodicity of twenty-three hours, fifty-six minutes, characteristic of sidereal objects, and the direction of the source at approximately eighteen hours right ascension. The front page of *The New York Times* of May 5, 1933, read "New Radio Waves Traced to Center of Milky Way." Interestingly, with both popular and scientific publicity, no university pursued the research. In 1936, G. W. Potapenko and D. F. Folland of the California Institute of Technology confirmed Jansky's results but were not funded for further research after their initial trial.

The next radio telescope was built by Grote Reber, a radio engineer in Chicago. During 1937 and 1938, Reber built a parabolic radio telescope in his backyard. With this new type of telescope, he identified radio sources in the constellations Cygnus, Cassiopeia, and Sagittarius, as well as signals from the Sun that had eluded earlier researchers. He also constructed a detailed radio map of the Milky Way.

World War II interrupted many scientists' research but advanced the technology of radio and the new field of radar. During the war, some scientists made near-simultaneous discoveries in radio and radar astronomy but for military and/or political reasons were unaware of one anothers' work. In 1942, J. S. Hey in England correlated radio noise with a large solar flare and identified the Sun as a radio source, but it was not announced until after the war. After 1945, benefits were gained from the technological developments such as the development of radar, high-frequency receivers, and antennas of large aperture and higher gain. Radio observations of the Sun and the galaxy were vastly improved, and the discovery of other radio sources began a race to build radio telescopes in the early 1950's. Radio astronomy, like its optical counterpart, became big science requiring funding from the federal government as well as underwriting by wealthy philanthropists.

Historically, radio astronomy was concerned with continuous emissions from various extraterrestrial sources. In 1951, however, a subdiscipline developed: the study of 21-centimeter line emissions. These emissions are the radio signatures of largely neutral hydrogen atoms first detected by Harold Ewen and Edward Mills Purcell of Harvard University. This technique of observing the 21-centimeter spectral line of neutral hydrogen provided a method of mapping the galaxy that was previously unclear to astronomers. It was discovered that hydrogen was a major component of interstellar space and the universe, accounting for about 75 percent of its mass.

During the 1950's, there was a rush to build radio telescopes and identify as many discrete sources as possible. Most radio sources coincided with optical sources such as galaxies and supernova remanents. The culmination of this data gathering produced new catalogs of the heavens and a reexamination of the more familiar optical sources. The discovery of interstellar molecules, radio galaxies, quasars, pulsars, and the solar wind have all influenced the direction of astronomy since that time. Radio astronomy has contributed significantly to theories of radio emissions, the origin of the universe, and the physics of energy mechanisms.

Richard C. Jones

FURTHER READING

Burke, Bernard F., and Francis Graham-Smith. *An Introduction to Radio Astronomy*. Cambridge, England: Cambridge University Press, 2002. A review of radio observations of stars, pulsars, quasars, radio galaxies, and the center of the Milky Way, including a detailed explanation of the cosmic microwave background. Technical.

Condon, J. J., and S. M. Ransom. *Essential Radio Astronomy*. Available online at http://www.cv.nrao.edu/course/astro534. Detailed notes from a National Radio Astronomy Observatory/University of Virginia course in radio astronomy. Includes links to chapters on pulsar physics and observations.

Hey, J. S. *The Evolution of Radio Astronomy*. New York: Science History Publications, 1973. A fine historical account of an emerging science from one of its pioneers.

Jansky, Karl G. "Electrical Disturbances Apparently of Extraterrestrial Origin." *Proceedings of the Institute of Radio Engineers* 21 (1933): 1387-1398. This paper is the primary source for radio astronomy and one of the most important for astronomy as a whole. Jansky describes many basic astronomical concepts as well as his announcement of an extraterrestrial source. Readable and informative.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who find standard introductory textbooks pitched at too low a level. Suitable for an audience with varied science and mathematical backgrounds. Covers all topics from solar-system objects to cosmology.

Lockman, F. J., F. D. Ghigo, and D. S. Balsar, eds. *But It Was Fun: The First Forty Years of Radio Astronomy at Green Bank*. Washington, D.C.: National Radio Astronomy Observatory, 2007. A collection of papers by key radio astronomers. Provides an observational and technological history of the development of radio astronomy with a focus on facilities at Green Bank Observatory in West Virginia. Also discusses the landmark discoveries of Jansky, Reber, and Bell.

Malphrus, Benjamin K. *The History of Radio Astronomy and the National Radio Astronomy Observatory: Evolution Toward Big Science*. Malabar, Fla.: Krieger, 1996. A history of the development of radio astronomy from the 1930's to the creation of the National Radio Astronomy Observatory. Accessible to the general reader, but also provides technical descriptions suitable for those with astronomy and electrical engineering backgrounds.

Sullivan, W. T., ed. *The Early Years of Radio Astronomy*. New York: Cambridge University Press, 1984. A fine historical account of the first fifty years of radio astronomy. The people and history of this emerging science are well documented, with fine historical photographs and illustrations.

Verschuur, Gerrit L. *The Invisible Universe: The Story of Radio Astronomy*. New York: Springer Praxis, 2006. Provides a history of developments in radio astronomy, and along the way describes the discovery of pulsars, quasars, and radio galaxies. Suitable for a general science course in college as well as for astronomy majors as background information.

See also: Archaeoastronomy; Coordinate Systems; Earth System Science; Extraterrestrial Life in the Solar System; Gravity Measurement; Hertzsprung-Russell Diagram; Infrared Astronomy; Milky Way; Neutrino Astronomy; Optical Astronomy; Solar Radio Emissions; Telescopes: Ground-Based; Telescopes: Space-Based; Ultraviolet Astronomy; X-Ray and Gamma-Ray Astronomy.

Red Dwarf Stars

Category: The Stellar Context

Red dwarf stars have the lowest mass and luminosity but largest population of stars in our galaxy. Developing theoretical models of the evolution of these stars and their planetary systems has expanded astronomers' understanding of stellar evolution and the circumstances under which planets might support life.

OVERVIEW

Red dwarf stars are low-mass, low-temperature, low-luminosity stars that occupy the lower, "tail" end of the main sequence on the Hertzsprung-Russell (HR) diagram. The majority are in spectral class M, although some authors also include late K stars (K5 onward) as well as some L-type stars (so-called ultracool dwarfs). Their red color is due to a low surface temperature, between approximately 2,100-3,800 kelvins, and they actually emit more energy in the infrared portion of the electromagnetic spectrum than in visible light. Therefore they are most easily found via infrared (IR) surveys, such as the Deep Near Infrared Survey (DENIS) and the 2-Micron All Sky Survey (2MASS) of the late 1990's.

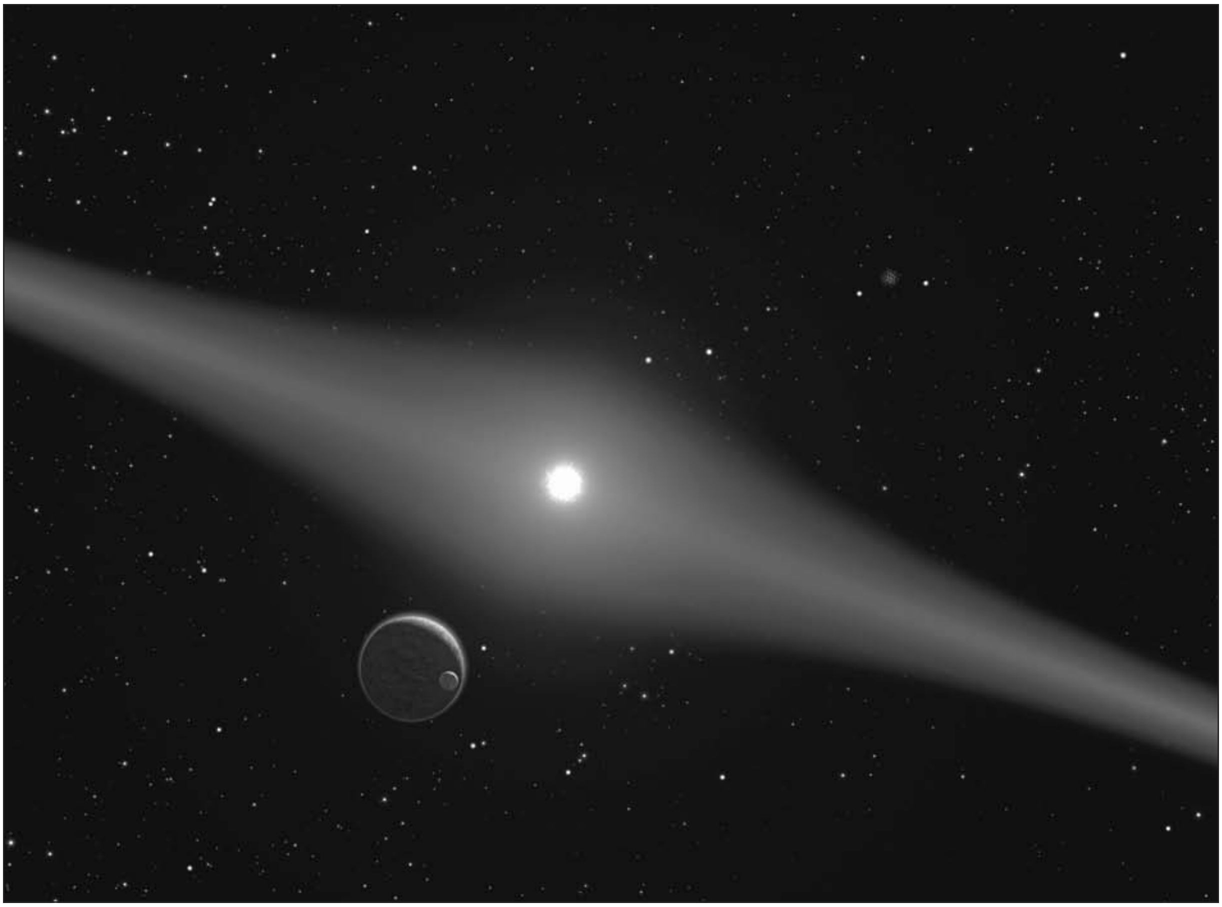
Because of their low surface temperature, the spectra of red dwarfs have distinctive absorption lines, especially titanium oxide (TiO), water, carbon monoxide, and vanadium oxide (VO). Their low temperature is caused by their low mass (0.08-0.6 times that of the Sun, a main sequence star), which limits their ability to generate energy through nuclear fusion. As a result, although red dwarfs are thought to be the most common type of star—comprising up to 80 percent of the total population of stars (and as much as half of total stellar mass of the Milky Way)—none is visible to the unaided eye. The brightest known red dwarf, AX Microscopium, has a disappointing 6.7 visual magnitude.

Not only do these stars have low luminosity (between 0.001 and 0.01 times that of the Sun), but their radii are also small, between 0.1 and 0.6 that of our Sun. The designation "red dwarf," therefore, is commonly used to prevent confusion between these "underachieving" stars and the impressively large "red giants" and supergiants that share a similar surface temperature (and hence color) with these stars. Given their low luminosity, it is not surprising that many red dwarfs remain undiscovered in our own celestial neighborhood. It is estimated that up to eight thousand red dwards are within 25 parsecs of the Sun, the majority of them undiscovered.

Among the nearby red dwarfs, two are of historical importance. Proxima Centauri is the third star in the Alpha Centauri multiple system and technically lies 10,000 astronomical units (AU) closer than its much brighter G and K class companions, making it the closest known star outside our solar system (4.22 light-years). Barnard's star, discovered by E. E. Barnard in 1916, has the largest proper motion of any known star, 10 arc seconds per year.

From the 1930's to the 1950's, astronomers struggled with models of energy generation and transportation in these stars. Because of their low mass, their central pressure and temperature are much less than in solar-type stars, leading to a significantly lower rate of hydrogen-to-helium fusion and hence less energy production. Another important difference is the mechanism used to transport energy from the core to the surface. In solar-type stars, radiative transfer is used outside the core until convection becomes more efficient in the outer quarter of the Sun. In red dwarf stars, convection occurs much deeper within the star, and in the coolest M dwarfs (M4 and later) the entire star is convective. This large-scale movement of material means that nearly 100 percent of the star's initial hydrogen is available as fuel over the course of the star's life (in contrast to an estimated 10 percent for the Sun). Greg Laughlin and his colleagues have computed the lifespans of these miserly stars to be hundreds to thousands of times greater than the current age of the universe: between 10^{11} and 10^{13} years. In their study of the future of the universe, Laughlin and Fred Adams have defined the end of the Stelliferous (star-filled) Age in which we now live by the death of the last generation of red dwarfs, predicted to be 10^{14} years after the universe's birth.

Many red dwarfs exhibit stellar activity reminiscent of the Sun, including spots, flares, and coronal mass ejections. UV Ceti or flare stars were discovered by Ejnar Hertzsprung, Adriaan van Maanen, and others in the 1920's, 1930's, and 1940's. Later shown to be red dwarfs, these variable stars suddenly brighten by several magnitudes and then decline over minutes to hours. For example, Proxima Centuri is a flare star. During these outbursts, a star can increase its ultraviolet and X-ray output by a hundred times or more, and some stars also exhibit radio outbursts. The general mechanism is similar to that of solar flares, although more energetic by a factor of a thousand or greater. BY Draconis red dwarf variables (named for the



An artist's conception of a terrestrial planet orbiting a red dwarf star. (NASA/ESA/G. Bacon, STScI)

multistar system BY Draconis) vary in visual brightness by only tenths to hundredths of magnitudes over days or months. This behavior is caused by the rotation of the star, carrying changing numbers of "starspots" and chromospheric areas of activity into view. For some stars, as much as 10-40 percent of the surface is covered in spots at any given time.

KNOWLEDGE GAINED

The study of red dwarf stars has added to our understanding of stellar structure and evolution, but even more important, it has caused astronomers to take a second look at long-held assumptions about planet formation and the definition of a "habitable" planet. Observations of young red dwarfs have found that the formation of protoplanetary disks (proplyds) is as likely around these stars as it is around higher-mass stars. For example, AU Microscopium, which lies 10 parsecs away and is about 12 million years old, has a disk of cold dust that shows signs of being in the late stages of planetary formation. However, because of the narrow definition of "habitable planet" (with its myopic view of solar-type systems), planets around red dwarfs were not considered objects of interest until the mid-1990's.

According to the most general definition, a planet is considered to be in a star's habitable zone if its surface temperature allows for the existence of liquid water. For a star of given temperature and luminosity, this designates a possible range of orbital distances. Given red dwarfs' low luminosity and surface temperature, any habitable zone would be narrow (about 0.02 to 0.2 AU) and lie very close to the star. This was thought to present insurmountable challenges for life. For example, it was assumed that planets close to a red dwarf would be tidally locked in such a way that they would keep one (exceedingly hot) face permanently turned toward the Sun and the other (brutally cold) face forever locked in darkness. Initial calculations suggested that on the cold side any atmosphere would be permanently frozen. In addition, astrobiologists were concerned about the effects of the stellar flares on life.

In 1994, the First International Conference on Circumstellar Habitable Zones revisited suppositions about habitable planets and found red dwarfs worth a second look. This led to the First Workshop on Habitability of M Star Planets, sponsored by the Astrobiology Institute of the National Aeronautics and Space Administration (NASA), in 2005. These conferences led to a dramatic rethinking of red dwarfs as hosts for life-sustaining planets. For example, more detailed calculations demonstrated that planets in close orbits might not be tidally locked, and even if they were, a planetary atmosphere with a pressure one-tenth that of Earth would have sufficient heat circulation to keep the atmosphere from freezing out, and if the pressure were similar to that of Earth (about 1,000 millibars), liquid water could be present. Development of an ozone layer could theoretically protect life from excessive ultraviolet exposure during flares. Flare activity, however, might not present any difficulty for the evolution of life around red dwarfs, since it has been shown that such activity decreases with the age of the star and lasts for a few million years for higher mass stars up to approximately a billion years for low-mass stars. Since these time frames are an insignificant fraction of the star's total lifespan, flare activity poses a serious threat to life only during the star's youth.

Even the presumed problem of the low luminosity (and hence small habitable zone) of red dwarfs was found to be resolvable. Because of red dwarfs' low mass, the way they die differs from that of other stars. Early-type M stars with a quarter solar mass or greater leave the main sequence to become red giants (as they exhibit a brief period of hydrogen-shell burning), but no red dwarfs will ever reach the necessary central core temperature to convert helium into carbon via the triple-alpha cycle. The surface temperature as well as the radius of a lower-mass red dwarf star actually increases near the end of its life, turning the red dwarf into a "yellow giant" and thereby increasing the size of the habitable zone. With surface temperature and luminosity comparable to those of the Sun, this yellow giant can survive for billions of years, yielding Sun-like conditions that could facilitate the evolution of life and the development of technologically advanced civilizations. Therefore, these unassuming low-mass stars may in fact represent ideal targets for Search for Extraterrestrial Intelligence (SETI) searches.

Prior to the mid-1990's, both theoretical investigations and observational searches for planets around red dwarfs were largely neglected. Beginning in 1998, with the discovery of red dwarf planets ranging in mass from several times that of Earth to masses similar to Neptune and Jupiter, the question of planet formation around low-mass stars became relevant. A major problem was the prediction that the core accretion method of forming large gas giant planets would not be feasible in red dwarf systems because material would accumulate too slowly to grow large planets before the gas in the protoplanetary disk would dissipate. A new model, proposed by Alan Boss, called the "disk instability mechanism," allows for the rapid creation of massive gas giants but predicts that red dwarf systems should tend to support fewer gas giant planets than more massive stellar systems. This prediction is borne out by current observations. Most prevalent appear to be "super Earths" and "failed Jupiters" (exposed high-mass cores that either never accumulated a gaseous envelop or had it stripped off by ultraviolet radiation).

CONTEXT

The discovery of red dwarf planets and theoretical considerations of their formation are ongoing areas of research, both of which add to our understanding of the wide variety of possible planetary systems and the methods by which different types of planets can be created and discovered. For example, because of the small masses of red dwarf stars and the small size of their habitable zones, habitable planets will cause greater (and more easily measurable) variations in radial-velocity surveys than in comparable solar-type systems. Transits of planets in front of their star are also more probable in red dwarf systems and will create greater variations in the star's apparent brightness. Microlensing experiments have also detected several planets around red dwarfs, including a possible 7 Jupiter-mass planet, the largest yet found orbiting a red dwarf. The orbital transit missions Corot (which launched on December 20, 2007) and Kepler (planned for launch in 2009), in addition to the upcoming Space Interferometry Mission (SIM) and proposed ground-based transit, microlensing, and radial velocity initiatives, are expected to increase the current knowledge of red dwarf planetary systems significantly.

Kristine Larsen

FURTHER READING

Adams, Fred, and Greg Laughlin. *The Five Ages of the Universe*. New York: Free Press, 1999. This book, aimed at a popular audience, summarizes the authors' research on the past and future of the universe. Highlights the importance of long-lived red dwarfs as defining the end of the age of stars.

Boss, Alan. "Rapid Formation of Super-Earths Around M Dwarf Stars." *Astrophysical Journal* 644 (2006): 179-182. This technical paper summarizes the properties of several red dwarf planetary systems and analyzes the successes and limitations of various methods of planet formation in describing these systems, including the authors' own original model.

Kaler, James B. *The Hundred Greatest Stars*. New York: Copernicus Books, 2002. Profiles one hundred famous stars, including several red dwarfs, for a popular audience. Both scientific and historical criteria are used to select the members of this stellar "hall of fame."

Percy, John. *Understanding Variable Stars*. Cambridge, England: Cambridge University Press, 2007. This undergraduate-level volume surveys the classification and properties of variable stars. The sections on UV Ceti (flare) stars and BY Draconis stars are relevant to the study of red dwarfs.

Reid, I. Neill, and Suzanne L. Hawley, eds. *New Light on Dark Stars: Red Dwarfs, Low-Mass Stars, and Brown Dwarfs*. Chichester, England: Springer Praxis, 2005. This technical volume explores the properties and observations of red dwarfs, brown dwarfs, and exoplanets. Although written for an expert audience, it is the standard work in this field and a vital reference for the study of red dwarf stars.

Scalo, John, et al. "M Stars as Targets for Terrestrial Exoplanet Searchers and Biosignature Detection." *Astrobiology* 7, no. 1 (2007): 85-166. This technical paper provides a valuable overview of the properties of red dwarf stars in the context of the potential habitability of their planets.

Tarter, Jill, et al. "A Reappraisal of the Habitability of Planets Around M Dwarf Stars." *Astrobiology* 7, no. 1 (2007): 30-65. This technical paper summarizes the conclusions of the 2005 NASA Astrobiology Institute Workshop on Habitability of M Star Planets. A valuable overview of the definition of habitable planet and the challenges and debates surrounding the likelihood of finding habitable planets around red dwarf stars.

See also: Brown Dwarfs; Extrasolar Planets; Gamma-Ray Bursters; Main Sequence Stars; Nemesis and Planet X; Novae, Bursters, and X-ray Sources; Nuclear Synthesis in Stars; Protostars; Pulsars; Red Giant Stars; Stellar Evolution; Supernovae; Thermonuclear Reactions in Stars; White and Black Dwarfs.

Red Giant Stars

Category: The Stellar Context

Red giant and supergiant stars are large stars with relatively cool surfaces that are in the last stages of their lives. Our planet Earth and the life on it are composed of chemical elements synthesized by nuclear fusion reactions occurring in such stars and ejected into space as the stars die.

OVERVIEW

Stars are not everlasting; they change through a series of stages analogous to the life cycle of a living organism. Stars are formed, or "born," from clouds of gas and dust called nebulae. They generate energy for most of their lives by fusing lighter atomic nuclei into heavier atomic nuclei in their interiors. They stop generating energy, or "die," when they exhaust their nuclear fuel. The red giant or red supergiant stage is a part of the stellar life cycle that a star goes through near the end of its energy-producing life.

Nebulae are large interstellar clouds of gas (mostly hydrogen, some helium, and traces of other elements) and dust grains, containing enough matter to form hundreds to thousands of stars. If the density in some region of the nebula is great enough, the gravitational pull between the particles will make the surrounding part of the nebula—perhaps about 10 trillion kilometers, or 1 light-year across—begin to contract. As the individual atoms are pulled inward, they gain speed and their energy increases, heating the gas. As the temperature rises, the contracting material begins to shine as an embryonic star, or protostar. The central part of the protostar gets hot enough that collisions between atoms occur with sufficient energy to strip away electrons from the atoms, and the gas (at least near the center) becomes ionized, consisting mostly of protons (bare hydrogen nuclei), much smaller amounts of other nuclei, and lots of free electrons.

When the core temperature of the protostar reaches a few million kelvins, the bare hydrogen nuclei (protons) move fast enough to overcome their electrical repulsion (known as the Coulomb barrier) and, via a mechanism known as "quantum mechanical tunneling," get close enough to fuse together. This first nuclear fusion process, through a series of intermediate steps, fuses four hydrogen nuclei into one helium nucleus, releasing energy. The four hydrogen nuclei have slightly more total mass than the one helium nucleus produced, and this slight excess in mass is converted into energy according to Einstein's famous equation $E = mc^2$, which states that energy, E, and mass, m, are related by a physical constant (the speed of light squared, or c^2). The nuclear energy produced as a result of this fusion stops the gravitational contraction of the protostar and turns it into a stable "main sequence" star, like our present Sun. The energy is radiated away from the star's surface as light and other forms of electromagnetic radiation at the same rate it is being produced by hydrogen fusion in the star's core.

Because a typical star is composed initially mostly of hydrogen, this main sequence stage of fusing hydrogen into helium in its core is the longest part of a star's energy-producing life. As hydrogen in the star's core is exhausted, the core, now mostly helium, begins to shrink again and grow hotter, as it did when it was a protostar. Hydrogen fusion is transferred to a hydrogen-rich shell surrounding the shrinking helium core. The star's outer layers expand, and the expanding stellar surface cools and turns red in color. (It is important to note the opposite but simultaneous processes that occur in the two regions: The star's core shrinks and heats up, while its outer layers expand and cool off.) The star becomes a red giant or red supergiant, so called because of its large size and red color.

The large surface area with accompanying low surface gravity results in an extremely powerful stellar wind, which blows gas outward. The low surface temperature allows dust grains to form, which are blown outward by the intense flow of photons. Both processes contribute to a significant loss of mass from the outer layers of red giants and supergiants. Perhaps as much as 10^{-4} solar masses are lost each year, forced outward at speeds of tens of kilometers per second.

Stars with about the mass of the Sun become red giants. The Sun is a star about midway through its main sequence stage. It formed about 4.5 billion years ago, and probably has another 3 to 6 billion years before it exhausts the hydrogen in its core and expands to become a red giant. It will grow until its diameter increases to about one hundred times its present size, when the Sun will engulf the orbit of Mercury and perhaps Venus. Even if the Sun does not expand enough to swallow Earth, the Sun will be about one thousand times brighter than it is now, and Earth's surface temperature will climb to between 1,000 and 2,000 kelvins.

Earth's atmosphere will escape into space, the oceans will boil away, and the surface will become a sea of hot, at least partly molten, rock.

The core of a red giant continues to shrink and heat until it reaches a temperature of about 100 million kelvins. At this temperature, the helium nuclei composing the core begin to fuse into carbon nuclei; three helium nuclei fuse into one carbon nucleus. As with hydrogen fusion, the total input mass of the three helium nuclei is slightly greater than the output mass of the carbon nucleus, and the mass difference is converted into energy. The high density of the contracting core has made the free electrons there degenerate, meaning they no longer behave as an ideal gas. As a result, helium fusion ignites explosively in a runaway process called the helium flash. However, this "explosion" in the core is not seen externally because the outer layers act like a blanket.

The energy released by the helium flash expands the core and lowers its temperature enough so helium fusion continues there at a more sedate rate. The outer layers shrink some and grow hotter, changing from red to orange or yellow. Although still a giant star, it is not as large as it was as a red giant. This is another period of stability, called the horizontal branch stage. It is somewhat like the main sequence stage, except now the star has two nuclear energy sources—helium fusion in its core and hydrogen fusion in a shell surrounding the core—and this stage is of much shorter duration than the main sequence stage, lasting only about one-tenth as long.

When the core has exhausted all its helium, it once again begins to shrink and heat up, trying to tap a new nuclear fuel. Helium fusion is transferred to a helium-rich shell surrounding the shrinking core, and hydrogen fusion continues in a hydrogen-rich shell still farther out. The outer layers again expand to accommodate the increased energy flow, and the stellar surface cools and once again turns red in color. The star becomes a red giant a second time, even brighter than before.

Stars like the Sun are not massive enough for their cores to gravitationally shrink enough to grow hot enough to initiate any more fusion reactions. A strong stellar wind and thermal pulsations puff off the outer layers as expanding bubbles of gas called planetary nebulae. (Planetary nebulae have nothing to do with planets. The name dates back to the 1700's, when, viewed through telescopes of that time, they looked round, like planets, and fuzzy, like nebulae.)

All that remains of the star itself is the exposed core, composed mostly of carbon nuclei and degenerate free electrons. This means the electrons are packed together as tightly as quantum mechanics allows. Consequently, the star cannot shrink further, so it cannot generate more energy by gravitational contraction to get hot enough to start new nuclear fusion reactions. The star shines only because it is very hot, but as it shines, it radiates its energy away and cools off. Such a star has about the mass of the Sun packed into a sphere about the size of the Earth, giving it an average density of approximately one metric ton per cubic centimeter. Since it was formerly the core of a red giant, at first it is very hot and shines with a white or bluish-white glow; then it is called a white dwarf. As it cools and fades, eventually it turns into a black dwarf.

When stars much more massive than the Sun exhaust the hydrogen in their cores, they become red supergiants. In stars exceeding at least eight times the Sun's mass, the core gravitationally contracts sufficiently to get hot enough to undergo a whole series of nuclear fusion reactions, each one proceeding more rapidly, that synthesize progressively heavier nuclei up to the element iron. Such a massive star quickly (on an astronomical timescale) develops an "onion-skin" interior, with an iron core surrounded by layers (like an onion) of lighter nuclei. The outer part of the star, still hydrogen-rich, has expanded to become so large that the star is called a supergiant, with a diameter hundreds and maybe up to a thousand times the diameter of our Sun.

Iron is the heaviest nucleus that can be formed in fusion reactions that release energy; to form heavier nuclei through fusion requires the input of energy. Since the iron core of a supergiant cannot tap any further fusion reactions to produce energy, it is unable to support itself against gravity and it collapses. The outer layers collapse too and rebound off the core, sending shock waves through the star. The star explodes as a core-collapse (or Type II) supernova, becoming billions of times more luminous than the Sun. The stellar explosion is so violent that energy is available to synthesize the elements heavier than iron, but only in relatively small amounts. These heavy elements, along with those formed in the star's interior before it exploded as a supernova, are dispersed into space, there to enrich the nebulae (clouds of gas) from which new stars and planetary systems will form. The atoms on Earth and in our bodies that are heavier than hydrogen and helium were made in massive supergiant stars that exploded as Type II supernovae before our solar system formed. A collapsed, compact remnant of the star itself may survive the explosion as either a neutron star or a black hole, depending on whether its mass is less than or greater than about two to three times the Sun's mass.

KNOWLEDGE GAINED

A well-known example of a red supergiant is Betelgeuse (Alpha Orionis). Anyone familiar with the night sky probably can pick out this star in the constellation of Orion the hunter. It appears as a bright red star in Orion's shoulder, one of the ten brightest-appearing stars in the night sky. Its red color tells us its surface is relatively cool, about 3,000 kelvins (only about half the Sun's surface temperature). It is located about 130 parsecs, or 425 light-years, away from Earth. Its distance together with its apparent brightness can be used to calculate its real brightness; it is about 40,000 times more luminous than the Sun. (That is why it appears so bright in the night sky, even though it is so far away.) The angular diameter of Betelgeuse has been measured to be almost 0.05 arc second, using a technique called speckle interferometry. Its angular size combined with its distance can be used to calculate its actual size, which is about six hundred to seven hundred times larger than the Sun; if placed at the center of our solar system, it would be about twice as big as the orbit of Mars and extend out into the asteroid belt.

Betelgeuse's large size makes its outer atmosphere unstable. Its brightness changes over timescales of weeks to years. The slower variations probably result from its outer layers expanding and shrinking, its size varying by up to 60 percent. The flickering observed over a few weeks is most likely caused by rising bubbles of very hot gases thousands of kilometers across. Poorly defined bright spots have been detected by speckle interferometry and by the Hubble Space Telescope; they may be these hypothesized hot bubbles, or possibly giant "storms" like those that occur in active regions on our Sun. The star has a strong stellar wind and is surrounded by a shell of gas and dust that has been blown outward to a distance of several thousand astronomical units. (For comparison, Earth is 1 astronomical unit and Neptune is 30 astronomical units from the Sun.)

The future of Betelgeuse is unclear. Its original mass may have been about 12 to 17 times the Sun's mass, but how much mass it has lost as a red supergiant is uncertain. It may be nearing the end of its life. If enough mass is expelled, Betelgeuse could become a white dwarf. If not, it will explode as a Type II supernova. Continued study of Betelgeuse may reveal more clues to the nature of this red supergiant and its future.

Above and to the right of Orion is the constellation of Taurus the bull. The bright orange star Aldebaran (Alpha Tauri), at one end of the V-shaped pattern of stars that makes up the bull's head, is an orange giant. (Orange giants are not quite as cool or as large as red giants, but Aldebaran is easy to spot not far from Orion and Betelgeuse.) Aldebaran is about 20 parsecs, or 65 light-years, away. It is about 370 times more luminous than and about 60 times larger than the Sun; if placed at the center of our solar system, it would fill more than half of Mercury's orbit around the Sun.

CONTEXT

Without the formation of heavier chemical elements by nuclear fusion reactions in the interiors of massive red supergiants, and the dispersal of these heavy elements into interstellar space when they explode as Type II supernovae, planets like Earth and life as we know it could not exist. The first stars and any accompanying planets to form in the early days of our Milky Way galaxy would have been composed of not much but hydrogen and helium. Planets similar to Jupiter could have formed, but not rocky/metallic planets like Earth.

The nebula from which the Sun and solar system formed about 4.5 billion years ago was composed about 2 percent (by mass) of atoms heavier than helium, having been enriched in these heavier elements by earlier massive stars that had exploded as Type II supernovae. This provided the raw materials for rocky planets like Earth and for life itself. The process of enrichment of interstellar gas continues today as massive stars quickly run through their life cycles and explode. The proportion of heavier elements in many nebulae now is up to about 5 or 6 percent.

Iona C. Baldridge

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A well-written college-level textbook for introductory astronomy courses. Provides an entire chapter on the post-main-sequence evolution of giant and supergiant stars, including a small section on Betelgeuse.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. One chapter covers the post-main-sequence evolution of giant and supergiant stars.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook in which much of one chapter deals with giant and supergiant stars and post-main-sequence stellar evolution. Thorough and well-written.

Jastrow, Robert. *Red Giants and White Dwarfs*. New York: W. W. Norton, 1979. A dated description of stellar evolution as it was understood in the late 1970's. The broad outline of stellar evolution presented here is still generally correct, though many of the details have changed. Simply written without mathematics, it can be read and understood easily. Several color plates illustrate concepts presented.

Sagan, Carl. *Cosmos*. New York: Random House, 1980. The companion volume to the popular television series. Although somewhat dated in its details, the overall picture it presents is still generally valid, and both novice and advanced readers will enjoy Sagan's way of making complicated topics understandable. Includes many color photographs and recommended readings.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A thorough college textbook for introductory astronomy courses, divided into many short sections on specific topics. Contains many sections on stellar evolution and giant and supergiant stars, including a description of Betelgeuse.

See also: Brown Dwarfs; Earth-Sun Relations; Gamma-Ray Bursters; Hertzsprung-Russell Diagram; Interstellar Clouds and the Interstellar Medium; Main Sequence Stars; Novae, Bursters, and X-Ray Sources; Nuclear Synthesis in Stars; Protostars; Pulsars; Red Dwarf Stars; Solar Evolution; Solar Magnetic Field; Stellar Evolution; Supernovae; Thermonuclear Reactions in Stars; White and Black Dwarfs.

S

Saturn's Atmosphere

Categories: Planets and Planetology; The Saturnian System

Data from Pioneer 11, Voyagers 1 and 2, and the Cassini spacecraft, combined with ground-based observations, show Saturn's atmosphere to be composed largely of hydrogen mixed with helium. Clouds of ammonia ice and other chemical components are sources for the various configurations on the planet's visible surface. This hydrogen-helium envelope likely covers a layer of metallic hydrogen that surrounds Saturn's rocky core.

OVERVIEW

Archaeologists have found recorded observations of Saturn dating from several hundred years B.C.E., inscribed in cuneiform on kiln-baked bricks. It was not until more than two millennia later that Galileo, using one of his early telescopes, was first able to see both the planet and its rings—though he did not recognize them as such. Original drawings made by Galileo indicate that he interpreted the rings as solid. At times he drew them as if they were open handles on a cup, and at other times as filled-in semicircles connected to the planet. In 1794, Sir William Herschel, carefully compared many weeks' worth of observations of subtle markings to deduce a planetary rotation period of 10 hours, 16 minutes, and 0.4 second. In the twentieth century it is realized that Saturn has a mean density of 690 kilograms per cubic meter, considerably less than that of water. Therefore it was realized that Saturn must be composed mainly of the lightest element, hydrogen. The presence of hydrogen was difficult to detect spectroscopically, but methane and ammonia were both observed in 1932.

Ground-based observations of Saturn have always been fruitful, as well as exciting, and continue to be so. However, astronomers' knowledge of the Saturn system was increased dramatically by four spacecraft encounters with the planet. The first was Pioneer 11, later called Pioneer Saturn, which made its closest approach to Saturn on September 1, 1979. It made many important observations of the entire Saturnian system but was concerned mainly with the nature of the surrounding radiation and particle environments. Images made by the photopolarimeter showed a butterscotch-colored planet with an indistinct pattern of subtly varying belts and zones. In other words, Saturn's atmosphere appeared much more subdued than Jupiter's.

More data were gathered during the first of the two Voyager encounters. Voyager 1 made its closest approach to Saturn on November 12, 1980. A few months later, its sister spacecraft, Voyager 2, made its closest approach on August 26, 1981. It was the outbound trajectory of the Voyager 2 spacecraft that was expected to provide good images of Saturn's southern hemisphere. Unfortunately, soon after closest approach, a problem arose with Voyager 2's scan platform, the movable platform upon which several of the instruments—including the two cameras—were mounted. Thus, most of the images obtained during the departure portion of the flyby show primarily the northern hemisphere.

Continuing analysis of all these observations has led to a reasonably consistent picture of Saturn's atmosphere, although much remains uncertain. One of the most obvious features of Saturn is its shape. It is flattened at its poles more than any of the other planets. Measurements estimated by the International Astronomical Union (IAU) in 1985 for Saturn's polar and equatorial radii were 53,543 and 60,000 kilometers, respectively.

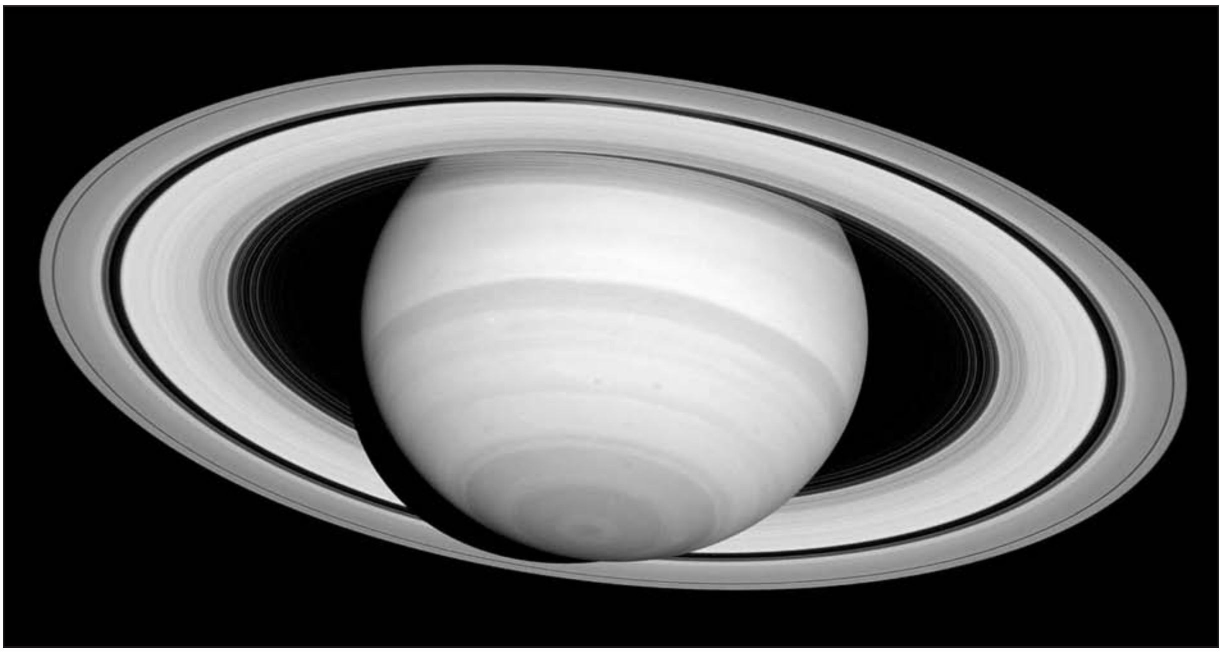
The shape of the planet, in association with measurements of other quantities, such as its gravitational field, composition, and heat production, is an important constraint on models of the planet's interior. The major chemical component of Saturn is hydrogen, but other elements are present as well. In particular, a heavy central core, composed of some sort of ice or rock, is proposed. To be consistent with gravity measurements, this core must contain between 10 and 20 percent of Saturn's mass. This proportion is considerably larger than the 2 or 3 percent that would be expected if Saturn had the same composition as the Sun. This central region may extend to about one-fifth of Saturn's radius.

Outside the core lies the hydrogen-helium atmosphere. In the visible part of the atmosphere, hydrogen is in its normal gaseous form, but deeper inside Saturn it is compressed by the weight of the overlying atmosphere. About halfway between the center and the visible surface, the pressure is so great, around 3 million bars, that hydrogen must change form and behave like a metallic fluid. For comparison, the atmospheric pressure at Earth's surface is about one bar.

It is this metallic region that could generate Saturn's magnetic field. Researchers believe that irregularities in this region produce slight asymmetries in the magnetic field, which, in turn, regulates the radio emissions observed by the Voyager spacecraft. It is for this reason that the observed period of 10 hours, 39 minutes, and 24 seconds (with an estimated error of 7 seconds) of this radiation is thought to represent the rotational period of the deep interior of the Saturn.

After hydrogen, the largest constituent of Saturn is helium, the second lightest element. The mass fraction of helium in Saturn has been estimated from infrared measurements to be 0.06 ± 0.05, significantly less than the value of 0.18 ± 0.04 found for Jupiter. This difference is thought to be caused by the fact that, at the lower temperatures found on Saturn, some of the helium becomes insoluble in the primarily hydrogen atmosphere. It is thought that it may form small droplets at some level and "rain out," dissolving again at a deeper, and hence warmer, level. This process would warm the atmosphere somewhat.

The observed average temperature of Saturn, derived from infrared measurements, is about 95 kelvins, compared with an equilibrium temperature of 82 kelvins, the temperature that would be expected if the planet were entirely dependent on solar radiation for heat. The implication is that Saturn radiates 1.78 ± 0.09 times as much heat as it receives. Some of this energy

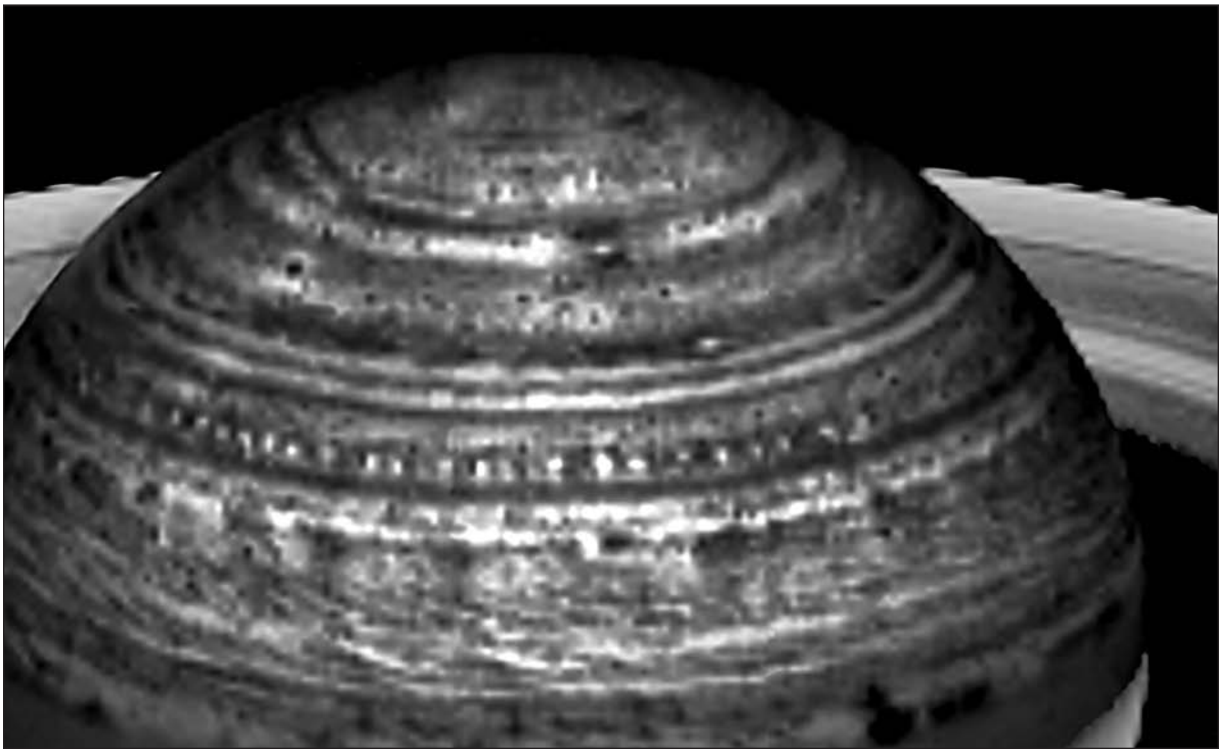


Saturn from the Hubble Space Telescope in 2004. (NASA/ESA/Erich Karkoschka, University of Arizona) is thought to be generated in the helium separation.

Because of its 27° orbital inclination, larger than the 23° inclination of Earth, Saturn has seasons. The period of the Voyager encounters corresponded to early spring in the northern hemisphere. The large heat capacity of Saturn's atmosphere, however, meant that the southern hemisphere, where it was early fall, would still be a few kelvins warmer than the northern hemisphere.

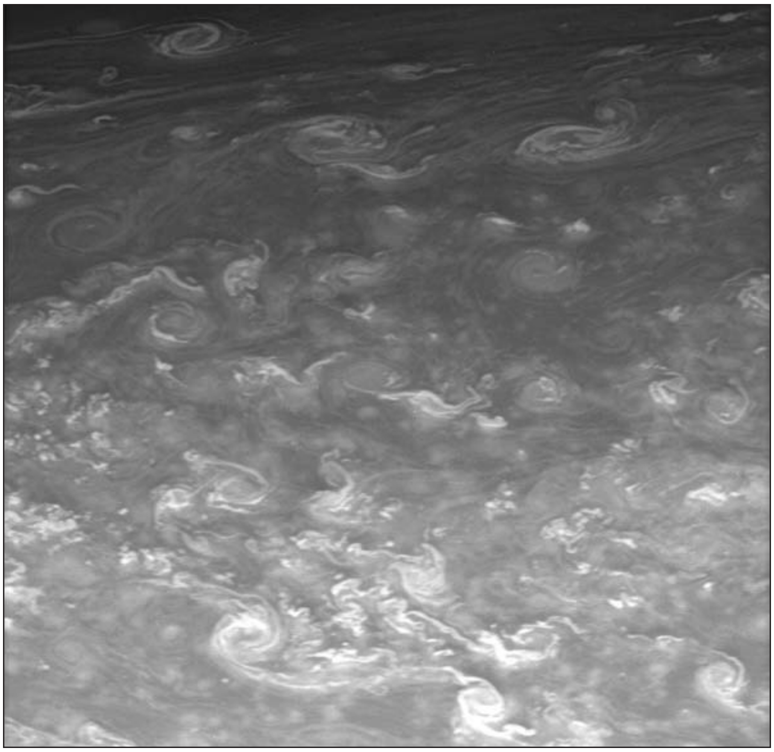
Although hydrogen and helium account for most of the mass of Saturn's atmosphere, they are not responsible for the patterns seen in the Voyager images. Despite the fact that only gaseous ammonia has been observed spectroscopically, these patterns are thought to be the tops of ammonia clouds. Theoretical models predict that below the ammonia clouds there are clouds of ammonia hydrosulfide and water ice. There is, however, little observational confirmation of these lower cloud layers. Lack of observational data makes it difficult to model the behavior and possible interactions of these multiple cloud layers. Whether they produce precipitation—ammonia or water precipitation in the form of either rain or snow—remains a matter of speculation.

The presence of ammonia clouds still fails to explain the appearance of the planet, since a cloud of small ammonia droplets would not create Saturn's observed butterscotch color. It should be remembered that the colors of many of the published spacecraft images are considerably enhanced (as "false-color" imagery) to make subtle variations distinguishable. Thus, an additional chemical component, termed a chromophore, has been postulated. Such a component must be capable both of existing at this level in the atmosphere and of providing the needed color. The complex radiation environment of Saturn's upper atmosphere could produce many types of chemicals. Among those suggested as cloud chromophores are various



Saturn's "string of pearls" formation can be seen at about 40° north latitude in this Cassini image of April 27, 2006. The bright "beads" are clearings in the clouds, which reveal the thermal glow of the planet below, and are regularly spaced, suggesting a large planetary wave. (NASA/JPL/University of Arizona) compounds of sulfur, phosphorus, and hydrazine, as well as various mixtures of organic compounds containing both hydrogen and carbon.

A comparison of two images of Saturn's atmosphere will show not only that there is an interesting variety of cloud shapes but also that they are all moving relative to one another. Most of the movement observed is longitudinal—that is, around the rotational axis. If the radio rotation period is used as the basic rotation rate, then the clouds show an alternating series of east-west jets. Close to the equator, these jets have a velocity of between 400 and 500 meters per second (the so-called equatorial super-rotation). The velocity decreases away from the equator until it reaches a retrograde speed of 25 meters per second at about 40° north and south latitudes. It then increases again to nearly 150 meters per second before decaying yet again. This pattern has been observed to 84° north latitude, always with posigrade jets moving considerably faster than the retrograde ones. The mechanism responsible for generating these jets is still uncertain. The motions in Earth's atmosphere are its response to an uneven solar heat input. The attempt to move warm air poleward produces, in the shallow (less than 1 percent of Earth's radius) rotating atmosphere, streams of high- and low-pressure areas encircling the midlatitudes. A similar mechanism could operate on Saturn, with observed surface motions decaying rapidly with height. An alternative suggestion is that Saturn's internal heat source produces a form of convection, and that the observed zonal flow around the rotational axis extends all the way through the planet, parallel to the rotational axis.



A close view of the far-north atmosphere, taken in May, 2008, from the Cassini orbiter, displays swirls and vortices in the hydrogen-helium "ocean." (NASA/JPL/Space Science Institute)

A number of additional features have been noted on Saturn's visible surface. In the southern hemisphere, there is a reddish spot which resembles a miniature version of Jupiter's Great Red Spot. Near the equator, there are various wispy clouds, which appear to be stretched by the strong wind shears present in this region. Farther north, at around 40° north latitude, there are three brown spots, each one a few thousand kilometers across. Just to the south of these is an area in which white blobs of cloud, possibly of convective origin, appear, evolve rapidly, and dissipate. In the mid-1990's, the Hubble Space Telescope detected a large white storm near the equator that eventually grew to nearly 20,000 kilometers in length before beginning to fade. This storm, referred to as Saturn's Great White Spot, is hardly unique. Large white storms of this type have been noted in astronomical records on a nearly thirty-year period going back well into the nineteenth century.

Just to the north of the brown spots, centered at about 47° north latitude, is the "ribbon feature," a fairly light band, about 5° wide, inside of which is a darker streak, around 1,000 kilometers wide, which threads an oscillatory north-south path along the center. The appearance of the individual peaks and troughs evolves rapidly over a few days, the average distance between adjacent peaks being 5,700 kilometers. This feature apparently represents some type of atmospheric wave. Measurements of the infrared emissions of this region show a large north-south temperature gradient, and it could be that this feature is the same type of atmospheric phenomenon as that which transfers equatorial heat northward on Earth.

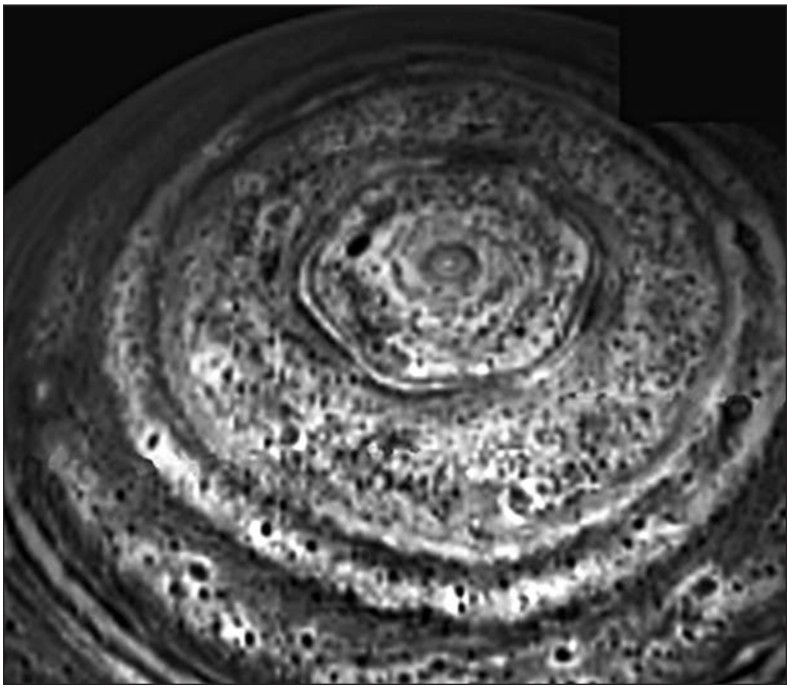
Despite the nearly equatorial trajectory of the Voyager spacecraft, some images of the north polar region were obtained. These show it to be populated with many small, fluffy clouds, close to which, at about 80° north latitude, was a regular hexagonal pattern, formed of long, thin, striated clouds that are moving at about 100 meters per second. The hexagonal pattern appeared to be stationary relative to the radio rotation period, with the clouds passing around its corners. The driving mechanism of this feature was a mystery, although the close association between its rotation rate and the radio rotation period could be significant. The Cassini spacecraft returned images in 2007 of Saturn's north polar region that again showed such a hexagonal pattern. This time the hexagon was nearly 24,000 kilometers across. Data suggested that the nearly perfect hexagonal pattern extended down into Saturn's clouds to a depth of nearly 100 kilometers. According to University of Oxford planetary scientist Leigh Fletcher, the appearance of this vortex was surprising. Apparently gas moved toward the pole and was compressed and heated as it dropped into the depths of Saturn's troposphere over the pole. The physical mechanism for this behavior remained unknown.

Cassini entered orbit about Saturn on July 1, 2004. During its primary four-year mission, the spacecraft returned impressive data about the ringed planet and its satellites, including the particles and fields environment surrounding Saturn. The spacecraft remained in nearly perfect health, and therefore it was funded for a two-year extended mission to continue the scientific harvest.

KNOWLEDGE GAINED

Spacecraft flybys of the Saturnian system and prolonged orbital observations by Cassini have dramatically increased astronomers' knowledge of Saturn's dynamic and complex atmosphere. Measurements of the hydrogen and helium abundance, the zonal circulation, the structure of the planet's gravitational and magnetic fields, and the derivation of the radio rotation rate would have been very difficult, if not impossible, to obtain from Earth and even from the Hubble Space Telescope.

These Cassini observations have enabled scientists to refine considerably their models of



The hexagonal shape at Saturn's north polar region (about 78° north latitude) was first discovered in the 1980's and seems to persist, appearing in this 2006 image from Cassini. (NASA/JPL/University of Arizona)

Saturn's internal composition and structure. This model envisages that the visible envelope, composed primarily of hydrogen and helium, extends about halfway to the planet's center; below this covering is a region of metallic hydrogen and a small rocky core. The deep atmosphere is warmed by Saturn's internal heat source and becomes cooler with distance from the core, reaching a minimum of about 85 kelvins at a pressure level of around 100 millibars at the tropopause. Above this level, solar heating becomes significant, and the temperature starts to increase again. Visible clouds are thought to consist of ammonia ice crystals with, possibly, ammonium hydrosulfide and water-ice cloud layers below them. Existence of these lower cloud layers would depend on the amounts of the various elements in this part of the atmosphere.

The Voyagers provided an extensive album of images of Saturn, documenting cloud motions and their morphologies. Such images, however, have increased scientists' understanding of the atmospheric dynamics only to a limited extent. Even the basic driving mechanism remains uncertain; it could be the planet's internal heat source or the Sun's external heat supply. The extent to which Voyager observations have increased scientists' knowledge can also be considered in terms of their understanding of atmospheres in general. The planetary atmosphere most comparable to Saturn's is that of Jupiter, which is slightly larger, rotates slightly faster and has a similar hydrogen and helium composition. Jupiter is, however, much closer to the Sun than Saturn, and it therefore receives several times as much solar energy.

Compared with Jupiter's highly turbulent atmosphere, Saturn's atmosphere, with its subtle belt-zone variations and widely separated spots, might appear to be relatively quiescent. Voyager observations, however, have shown that this is far from the case. Saturn's atmosphere moves at velocities up to 500 meters per second, compared with about 200 meters per second on Jupiter. Jupiter observations also suggested that Saturn's alternating belt-zone pattern was associated with alternating eastward and westward jets. Saturn observations, which had no such obvious correlation, showed that this model was overly simplistic. Another interesting difference between the two planets is that on Jupiter westward jets reach much the same speed, relative to the radio rotation period, as eastward jets. The situation on Saturn is far less symmetrical, with eastward jets dominating.

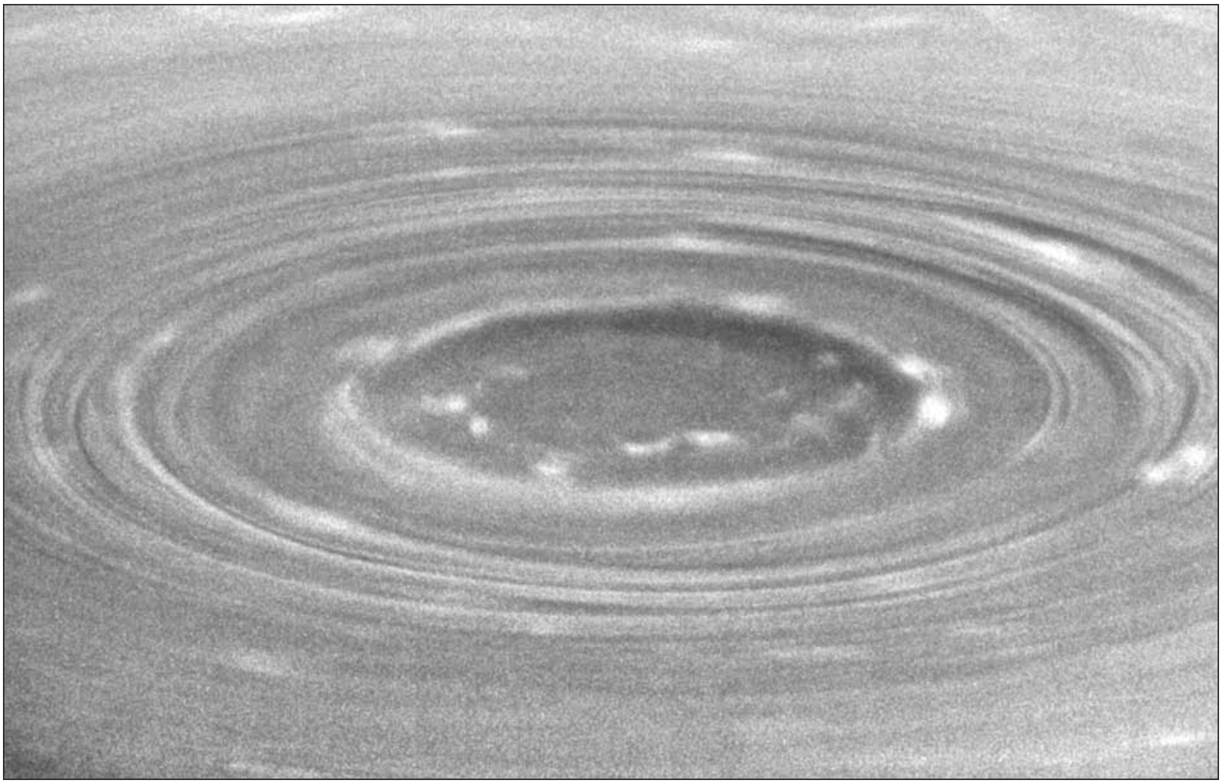
Circulation patterns in Earth's atmosphere are still not fully understood. Thus it is not surprising that scientists' knowledge of the atmospheres of the outer planets remains fairly basic. The measurements of the Voyager and Cassini spacecraft contribute to that knowledge mainly by increasing the number of observations upon which models can be based.

One such Cassini observation was referred to by Internet-based National Aeronautics and Space Administration (NASA) reports as Saturn "riding the wave." The wave pattern in the atmosphere is visible from Earth only every fifteen years. Earth has a similar oscillation, but its period is only two years. Jupiter has a similar oscillation, but with a four-year period. Saturn's oscillation in question was under observation with ground-based telescopes and Cassini's Composite Infrared Spectrometer. This oscillation involves temperatures in Saturn's upper atmosphere switching with altitude in a hot-cold pattern that, when graphed in three dimensions, assumes a shape much like the stripes that wrap around a candy cane. These temperature oscillations result in winds changing direction from east to west and back again, and it is this pattern to which scientists were referring as the wave pattern discovered on Saturn.

Beginning in December, 2007, Cassini imaged a storm in Saturn's southern hemisphere that produced lightning discharges with energies more than ten thousand times that of typical lightning on Earth. Cassini's radio and plasma wave instruments actually picked up the lightning about a week before the storm itself could be identified by the spacecraft's imaging cameras. Scientists then used the periodic appearance of the storm to confirm Saturn's rotation rate.

CONTEXT

If it had not been for the Pioneer, Voyager, and Cassini missions, knowledge of Jupiter and



At Saturn's south polar vortex in this July, 2008, Cassini image, bright clouds can be seen within the inner ring, providing tantalizing clues to how heat energy may move through Saturn's atmosphere. (NASA/JPL/Space Science Institute)

Saturn would be limited to that obtainable from Earth or from Earth-orbiting observatories such as the Hubble Space Telescope. Although probes have not entered Saturn's atmosphere, they have still provided many observations that contribute to astronomers' understanding of it. For the first time, researchers could see how the planet changes in appearance when it is viewed from different directions. Monitoring of each spacecraft's radio signal as it disappeared behind Saturn provided information about the local atmospheric temperature profile (the variation of temperature with height). Proximity of these spacecraft to the planet also allowed study of the structure of individual features.

Prior to these dedicated spacecraft missions, Saturn was a greatly appreciated but little understood object. Only its most global properties had been investigated. It was believed to be a typical "gas giant"—a smaller, colder, and less colorful version of Jupiter. Voyager observations revealed this concept to be only partly true. The elemental abundances derived for Saturn are different from those of Jupiter, the most striking difference being the hydrogen-to-helium ratio, which, combined with the planet's unexpectedly high temperature, led to the idea that its helium becomes depleted by becoming insoluble and precipitates into the lower atmosphere.

Very little was known about Saturn's atmospheric dynamics before Voyager images were obtained. Tracking motions of the occasional atmospheric spots visible from Earth had suggested a predominantly zonal flow, strongest near the equator. It was only with the Voyager flybys that periodic radio emissions from the planet were observed. The association of these with rotation of Saturn's interior provided a plausible base velocity against which the motion of other features could be measured. Prior to this finding, velocities could only be given relative to an arbitrary feature.

The Pioneer 11 encounter provided some intriguing information about Saturn's atmosphere; the two Voyager encounters supplied far more. All these flyby missions, however, could provide only a brief look at the planet. Ground-based observations suggest that the large-scale zonal flow is probably fairly stable. The appearance of the individual belts and zones can change, however, while smaller individual features can evolve quite rapidly. Some of these changes may be seasonal effects, while others could be caused by various types of instabilities in the same way that the fairly constant solar heating of Earth produces a perpetual sequence of low- and high-pressure regions.

To investigate these effects further, it was necessary to observe the planet over a longer period than is possible with a flyby mission. The Hubble Space Telescope was deployed from the space shuttle *Discovery* in April, 1990. After a troubled start requiring repair missions, Hubble was able to begin taking unprecedented images, including on occasion some of Saturn. Although Hubble could capture images only at a much lower resolution than the Voyager spacecraft had achieved, Hubble did have the capacity to observe in spectral bands that could not be seen from Earth because of absorption by the atmosphere.

The obvious next step in Saturn investigation was to place an orbiter about the ringed planet, one that could use encounters with its many satellites in addition to propulsion system firings to alter its course so that the spacecraft could swing close to interesting objects; a principal focus on studying the large satellite Titan and its atmosphere was given to this follow-on mission. The National Aeronautics and Space Administration and the European Space Agency named the follow-up probe to the Voyager flyby missions Cassini after the astronomer who discovered several of Saturn's satellites. Cassini was launched on October 15, 1997, and after several gravity assists in the inner solar system, it flew past Jupiter and then was redirected to enter the Saturnian system and conduct its primary science mission. That primary mission ended in 2008, but as the spacecraft was still in fully functional condition, a two-year extended mission was conceived and funded.

Cassini found confounding aspects in Saturn's dynamic and complex atmospheric structure. White spots, which were also seen with the Hubble Space Telescope, were seen to develop and diminish. However, perhaps the most perplexing atmospheric features were a hexagonal pattern about the north pole 25,000 kilometers across and another similar feature found later to exist around the south pole. The nature and stability of features such as these polar ones remains under investigation.

David Godfrey

FURTHER READING

Alexander, Arthur Francis O'Donel. *The Planet Saturn: A History of Observation, Theory, and Discovery*. New York: Macmillan, 1962. Alexander catalogs the historic observations of Saturn through the ages until approximately the middle of the twentieth century. For anyone interested in the history of Saturnian observations, this is an essential resource. The way the subject matter is divided into very small sections, however, makes some parts of the book difficult to read.

Benton, Julius. *Saturn and How to Observe It*. New York: Springer, 2005. An observer's guide to astronomy focusing on Saturn and what can be seen with modest telescopes. Provides plenty of information about the Saturnian system for the amateur astronomy buff.

Bortolotti, Dan. *Exploring Saturn*. New York: Firefly Books, 2003. A look at the Cassini-Huygens mission for a younger audience. Full of charts, photographs, a section on observing Saturn, and a discussion of the historical development of Saturn studies, from antiquity to the launch of Cassini.

Consolmagno, Guy. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994. A text aimed at both college science and nonscience majors alike. Presents subjects using low-level mathematics but uses integral calculus where required. Demonstrates how the area of planetary science progresses by questioning previous understandings in light of new observations.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system, from early telescopic observations through the space missions that had investigated all planets (with the exception of Pluto) by the publication date. Takes an astrophysical approach to place our solar system in a wider context as just one member of similar systems throughout the universe.

European Space Agency. *The Atmospheres of Saturn and Titan*. ESA SP-241. Paris: Author, 1985. This report of the proceedings of a conference held in 1985 covers the then-current understanding of Saturn and its largest moon, Titan, and anticipated the Cassini mission to the two bodies, launched in October, 1997.

Harland, David M. *Cassini at Saturn: Huygens Results*. New York: Springer, 2007. The text provides a thorough explanation of the entire Cassini program, including the Huygens landing on Saturn's largest satellite. Essentially a complete collection of NASA releases from the start of Cassini flight operations through the majority of Cassini's seventy orbits of its primary mission. Cassini's primary mission concluded a year after this book entered print. Technical, but accessible to a wider audience.

_______. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002. A technical description of the Cassini program, its science goals, and the instruments used to accomplish those goals. Provides a historical review of pre-Cassini knowledge of the Saturn system. Written before Cassini arrived at Saturn.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. The chapter on Saturn covers the Saturnian system and spacecraft exploration of it.

Hunt, Garry E., and Patrick Moore. *Atlas of Saturn*. London: Mitchell Beazley, 1982. This book, which is accessible to anyone with an interest in the planets, summarizes knowledge of Saturn both before and after the Pioneer and Voyager encounters. Provides detailed descriptions of the Voyager observations, as well as the spacecraft and the instruments that made them.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, magnetic fields, and atmospheres. Filled with figures and photographs. Accessible to the serious general audience.

Lovett, Laura, Joan Harvath, and Jeff Cuzzi. *Saturn: A New View*. New York: Harry N. Abrams, 2006. A coffee-table book replete with about 150 of the best images returned by the Cassini mission to Saturn. Covers the planet, its many satellites, and the complex ring systems.

Morrison, David. *Voyages to Saturn*. NASA SP-451. Washington, D.C.: Government Printing Office, 1982. Written by a member of the Voyager imaging team, this account of the Voyager missions to the Saturn system describes the people and events that contributed to the encounters and how they changed scientists' understanding of the planet. Includes many of the beautiful color pictures produced from the Voyager imaging data.

Russell, Christopher T. *The Cassini-Huygens Mission: Orbiter Remote Sensing Investigations*. New York: Springer, 2006. Provides a thorough explanation of the remote-sensing investigations of the Cassini orbiter and the Huygens lander. Outlines the scientific objectives of all instruments on the spacecraft and describes the planned forty-four encounters with Saturn's moon Titan. Only the science returns to 2006 are covered.

See also: Auroras; Brown Dwarfs; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Jovian Planets; Jupiter's Interior; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits: Couplings and Resonances; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Saturn's Interior; Saturn's Magnetic Field; Saturn's Ring System; Saturn's Satellites.

Saturn’s Interior

Categories: Planets and Planetology; The Saturnian System

Saturn, the second largest planet in the solar system, is famous for its ring system, but the key to understanding the planet is in understanding its interior structure. Though it is similar to Jupiter and likely shares a similar origin, Saturn is not just a smaller version of Jupiter. While Saturn is called a gas giant planet, its interior is more than just a big ball of gas.

OVERVIEW

Jupiter and Saturn are the two largest planets in the solar system. Both planets are believed to have formed directly out of the accretion disk of material surrounding the Sun during its formation. As a result, both planets would be expected to have similar compositions and structures. There are important differences between Jupiter and Saturn, however. Because Saturn formed directly from the material accreting to form the Sun, it is expected to have a composition similar to that of the Sun and Jupiter. Indeed, Saturn is composed primarily of hydrogen and helium, the main constituents of Jupiter and the Sun.

Saturn is the second largest planet in the solar system, having about 30 percent of Jupiter’s mass, yet that mass is spread over a volume that is about 60 percent that of Jupiter. This gives Saturn a much lower density than that of any other planet in the solar system. One reason for this low density is that hydrogen, the main component of Saturn and Jupiter, is very compressible. Saturn’s lower gravity, due to its lower mass, compresses its hydrogen less than does Jupiter. Thus, Saturn is much more distended than Jupiter.

Though Saturn’s primary composition is hydrogen and helium, it has also accumulated heavier elements through collisions with smaller bodies since it formed. Some of the material is iron and silicates (rocks), and these materials sink to the center of the planet to form a core. Such a core would be under extreme pressure and temperature, with the temperature of the core is probably near 10,000 kelvins. Since Saturn is in the outer solar system, many of the bodies colliding with it would also contain ices—not only water ice but also frozen carbon dioxide, frozen methane, and frozen ammonia. These ices, heavier than hydrogen, also sink deep into the planet. However, at the pressure and temperature at that depth, the ices are in a liquid state. The term “liquid ices” is often used to describe the state of these fluids.

Saturn has an oblateness of 0.98 (its equatorial diameter is 9.8 percent greater than the pole-to-pole diameter), making Saturn the most oblate of the solar system’s planets. Though this level of oblateness constrains the possible size of Saturn’s core, there is some disagreement among planetary scientists as to the size of Saturn’s core. Many models suggest that Saturn has a core that is about 10 to 15 times the mass of Earth. This is similar in mass to Jupiter’s core, but it is less compressed, having a diameter perhaps in excess of 20,000 kilometers. Though Saturn’s core is about the size of Jupiter’s core, it is a much larger percentage of Saturn than Jupiter’s core is of Jupiter. A layer of liquid ices of perhaps up to 10,000 kilometers thick likely sits on top of the core.

The vast majority of the rest of Saturn is composed of hydrogen and helium. These are in a gaseous state in the outermost parts of the planet. Clouds of ammonia ice and other ices exist in the upper few hundred kilometers. Below about 2,000 kilometers beneath the clouds, the pressure is so high that hydrogen is compressed into a liquid state. However, the temperature and pressure of Saturn’s interior are well beyond the critical point of hydrogen. The critical point, in chemistry, is the temperature and pressure at which there is no clear distinction between liquid and gas fluid states. Therefore, there is no clear boundary between the gaseous upper portion of Saturn and the liquid hydrogen inner portion. As depth increases, the hydrogen gradually becomes more and more liquid-like. The bulk of Saturn is in this liquid state, even though it is called a “gas” giant planet.

Hydrogen begins to exhibit typical metallic properties, such as electrical conductivity, when it is under sufficient pressure and temperature.

Saturn Compared with Earth

Parameter	*Saturn*	*Earth*
Mass (10^{24} kg)	568.46	5.9742
Volume (10^{10} km^3)	82,713	108.321
Equatorial radius (km)	60,268	6,378.1
Ellipticity (oblateness)	0.09796	0.00335
Mean density (kg/m^3)	687	5,515
Surface gravity (m/s^2)	10.44	9.80
Surface temperature (Celsius)	–160	–88 to +48
Satellites	60	1
Mean distance from Sun millions of km (miles)	1,434 (884)	150 (93)
Rotational period (hrs)	10.656	23.93
Orbital period (days)	10,760	365.25

Source: National Space Science Data Center, NASA/Goddard Space Flight Center.

On Saturn, these conditions are believed to be met near 25,000 kilometers beneath the cloud layers. Thus, from the top of the liquid ices layer out to nearly 55 percent of Saturn's radius, the hydrogen is in this liquid metallic state. Jupiter also has a liquid metallic hydrogen mantle, but on that planet this layer is far larger than on Saturn. As with Jupiter, the magnetohydrodynamics of Saturn's liquid metallic hydrogen is the source of the planet's magnetic field. Planetary magnetic fields are believed to be produced through the motion of a highly conducting fluid through a magnetic field. The conductor then reinforces the existing field, creating a rather stable and permanent planetary field. This is the dynamo model of planetary magnetic fields. Because Saturn's liquid metallic hydrogen layer is far smaller than Jupiter's, Saturn's magnetic field is much weaker than Jupiter's, being only about 3 percent as strong. Even so, it is still nearly six hundred times stronger than Earth's magnetic field.

Saturn, like Jupiter, radiates more energy than it gets from the Sun, almost 2.9 times what it gets from the Sun. Jupiter radiates energy through kelvin-Helmholtz contraction, a form of hydrostatic contraction that permits a fluid body to compress, generating thermal energy. Saturn has much less mass than Jupiter. Theory indicates that kelvin-Helmholtz contraction is unable to produce the level of thermal energy needed to account for observations of Saturn. However, Saturn's temperature is lower than Jupiter's. At the temperature and pressure of Saturn's liquid hydrogen layers, helium, another major component of Saturn, precipitates out. Droplets of helium, heavier than hydrogen, sink toward the planet's deeper interior. The sinking helium releases gravitational energy in the form of heat, warming the planet. This thermal energy is nearly double the energy that the planet gets from the Sun, resulting in Saturn's observed emissions.

KNOWLEDGE GAINED

Most observations of Saturn have, of necessity, been made from Earth. The interior of the planet, however, cannot be studied from afar. Spacecraft sent to Saturn have yielded more data, but there is still considerable debate about the nature of Saturn's interior. Mathematical models of Saturn's interior are made based on observed characteristics, but more research is needed. The exact nature of Saturn's interior is therefore still the object of much speculation.

Because Saturn is shrouded in clouds, astronomers are unable to measure Saturn's rotational rate directly using observations from Earth. The planet's magnetic field should rotate with the planet. In 1981, Voyager 2 measured a rotational rate of a bit over 10 hours, 39 minutes. In 2007, however, scientists at the National Aeronautics and Space Administration's (NASA's) Jet Propulsion Laboratory released a new finding of 10 hours, 32 minutes, and 35 seconds for Saturn's rotational rate using data from the Cassini orbiter. The discrepancy between these measurements may be explained by the fact that it is unusually difficult to determine the rotational rate for Saturn because the planet's magnetic axis is nearly the same as its rotational axis, providing little change in the magnetic field for spacecraft to measure as the planet rotates.

Saturn and Jupiter both likely formed from the cloud of gas swirling together to form the Sun. Thus, these two gas giants would be expected to have very similar compositions to the Sun. Indeed, like the Sun, Jupiter and Saturn are both composed mostly of hydrogen and helium. Saturn's atmosphere, however, contains quite a bit less helium than either the Sun or Jupiter. This finding was a mystery until astronomers found that Saturn radiates far more energy than it gets from the Sun. The explanation that helium precipitating to lower levels of Saturn could heat the planet also explains why the upper portions of Saturn, which can be observed, are deficient in helium.

Hydrogen and helium are the primary constituents of Saturn. Both Jupiter and Saturn are somewhat enriched in elements heavier than hydrogen and helium when compared with the Sun. This is expected because both planets have, over the several billion years since the planets' formation, accreted planetesimals, asteroids, and comets. Saturn is considerably more enriched in these heavier elements than is Jupiter, however, and the reason is unclear. Saturn may have simply accreted a greater percentage of these bodies, or it may have collected less hydrogen and helium when it formed. Further research is needed to answer the question of why Saturn and Jupiter have this compositional difference.

Studies of Saturn and Jupiter have led some astronomers to theorize that these planets may have formed somewhat closer together than they currently are in the solar system. Jupiter migrated closer to the Sun, and Saturn migrated somewhat farther from the Sun, to its current position.

CONTEXT

The two largest planets in the solar system, Jupiter and Saturn, probably formed in the same manner, at about the same time, in the same part of the disk of material swirling together to form the Sun. Thus, they would be expected to be very similar. Indeed, they are similar, but there are important differences between them. Understanding those differences will cast light on the conditions under which gas giant planets form. In turn, understanding the formation of the gas giants will improve our understanding of the formation of other planets in the solar system, such as Earth, as well as the planetary systems of other stars.

The planet Saturn, because of its great distance from Earth—more than 1.2 billion kilometers at its closest—has mostly been studied via telescopes here on Earth. However, four spacecraft have investigated Saturn up close. The first was Pioneer 11, which passed closest to Saturn on September 1, 1979. Then, two Voyager spacecraft visited Saturn, with Voyager 1 passing closest to Saturn on November 12, 1980, and Voyager 2 flying past Saturn on August 25, 1981. No spacecraft visited Saturn until more than two decades later, when the Cassini orbiter entered orbit around Saturn on July 1, 2004. Cassinni has been studying Saturn and its satellites ever since. Although Cassini carried the Huygens probe to study Saturn's satellite Titan, no atmospheric probe was carried to study Saturn's atmosphere or interior, so all studies of Saturn's interior must be made by inference from observations of Saturn's exterior and of the planet's magnetic field. This means that scientists do not yet have a firm grasp of Saturn's interior, and further research is needed to understand this planet's interior structure fully.

Raymond D. Benge, Jr.

FURTHER READING

Bortolotti, Dan. *Exploring Saturn*. New York: Firefly Books, 2003. A look at the Cassini-Huygens mission for a younger audience. Full of charts, photographs, a section on observing Saturn, and discussion of the history of our knowledge of the Saturnian system from antiquity to the launch of Cassini.

Corfield, Richard. *Lives of the Planets: A Natural History of the Solar System*. New York: Basic Books, 2007. A history of planetary exploration that emphasizes the exploration more than the planets themselves.

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001. An upper-division or graduate-level textbook on planetary sciences, with a short description of the interior structures of gas giant planets. This text requires a fairly high level of mathematical familiarity.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. An excellent college-level introductory astronomy textbook. An entire chapter is devoted to Jupiter and Saturn.

Harland, David M. *Cassini at Saturn: Huygens Results*. Chichester, England: Praxis, 2007. This book is mostly about Saturn's moons, but it also gives a good description of data from the Cassini orbiter.

Irwin, Patrick. *Giant Planets of the Solar System: An Introduction*. New York: Springer, 2006. An overview of all four gas giants in the solar system, this text is written at a level for advanced students. It covers all aspects of the planets, including theories of formation, and has a very good bibliography.

Lovett, Laura, Joan Harvath, and Jeff Cuzzi. *Saturn: A New View*. New York: Harry N. Abrams, 2006. A coffee-table book with about 150 of the best images returned by the Cassini mission to Saturn. Covers the planet, its many satellites, and the complex ring system.

Morrison, David. *Voyages to Saturn*. NASA SP-451. Washington, D.C.: Government Printing Office, 1982. A good overview of the Pioneer and Voyager missions to Saturn, and some of their initial findings, with numerous illustrations. Some of the information is dated, but still useful.

See also: Auroras; Brown Dwarfs; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Jovian Planets; Jupiter's Interior; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits: Couplings and Resonances; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Saturn's Atmosphere; Saturn's Magnetic Field; Saturn's Ring System; Saturn's Satellites.

Saturn's Magnetic Field

Categories: Planets and Planetology; The Saturnian System

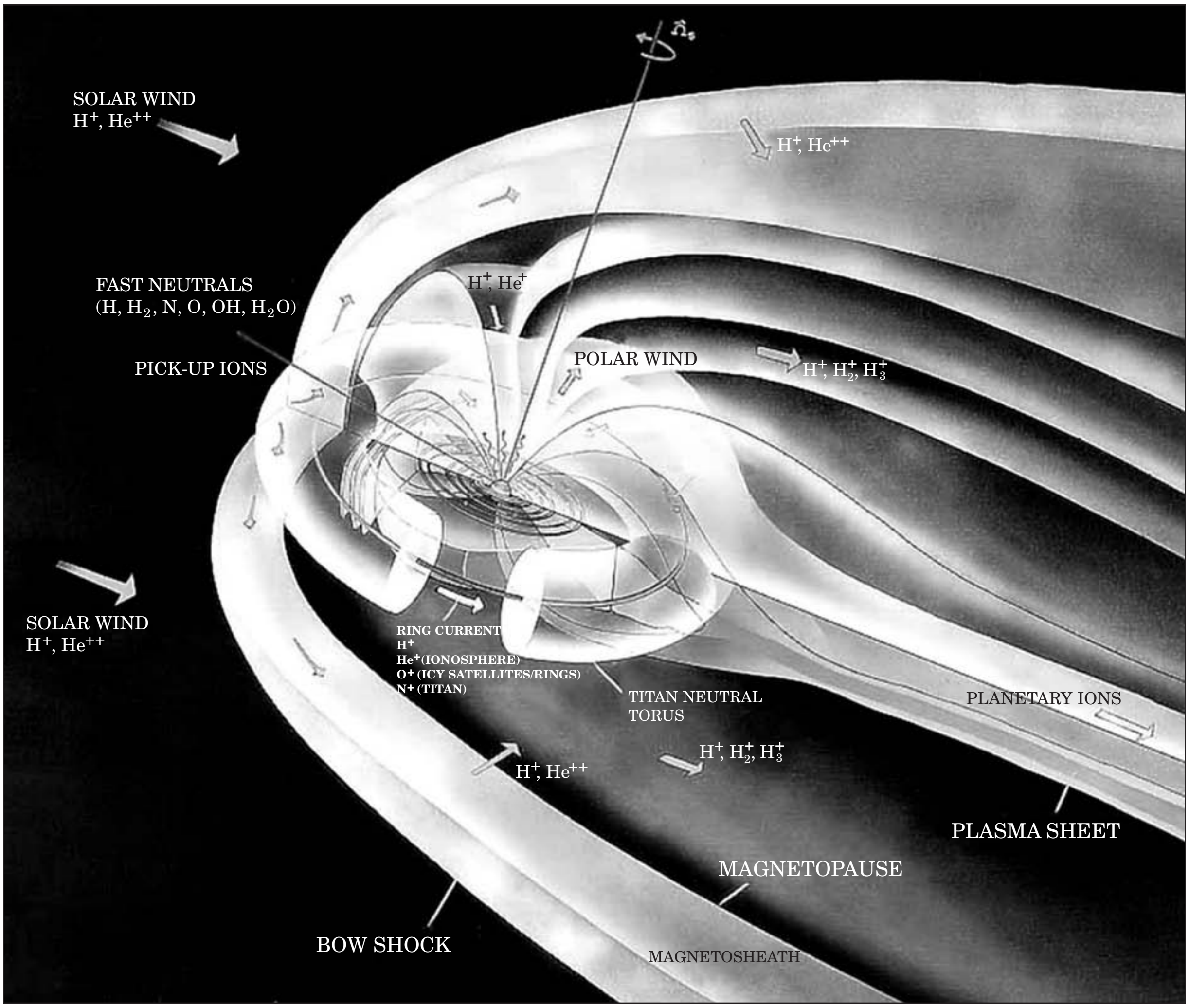
The magnetic field of Saturn, which was discovered and first analyzed from data collected by Pioneer 11, Voyager 1, and Voyager 2, reveals similarities to the magnetic fields of both Earth and Jupiter yet has many distinctive features. The Saturnian magnetosphere provides yet another cosmic laboratory for the study of astrophysically important processes such as collision-free shocks, charge-accelerating processes, plasma-wave modes, particle trapping, diffusion, and a host of related phenomena. The Cassini spacecraft provided an orbital platform from which those processes were studied on a regular basis for many years beginning in 2004.

OVERVIEW

After it was established in the early 1960's that decimetric radio emissions of Jupiter, discovered in 1955, were emitted by high-energy electrons trapped in the giant planet's intense magnetic field, a search for the scaled-down version of similar radiation from Saturn led to observation of radio noise from Saturn. Subsequently, a first model of Saturn's magnetic field was developed. Cassini observations resulted in the greater fidelity of that model and revealed that Saturn's satellite Enceladus possesses its own magnetic field and emits radio waves of a complex nature.

Initial concrete evidence for the existence of the Saturnian magnetic field was provided by Pioneer 11's magnetometer. The bulk of Saturn's magnetic field is believed to be generated by rapid internal motion of the metallic hydrogen that surrounds the planet's rocky central core, forming a dynamo and resulting in a field that resembles that of Earth and Jupiter. Field strength at cloud-top level over Saturn's equator was found to be 0.2 gauss, roughly a third of the equatorial geomagnetic field. By comparison, Jupiter's magnetic field at cloud-top level is ten times stronger than that of Earth. Saturn's magnetic axis and rotational axis are nearly coincident. The field is believed to originate at a greater depth than do those of Earth and Jupiter in relation to their respective radii. Moreover, the rotational and magnetic axes of both Earth and Jupiter make a nearly 11° angle with their respective rotational axes. Pioneer 11 and Voyager data indicated that the center of Saturn's magnetic dipole axis is offset by 4 percent of Saturn's radius to the north of the center of the planet, and the polarity of the field is reversed with respect to Earth's polarity. These facts combine to give Saturn's external magnetic field a perfect symmetry, with none of the wobbles that characterize the fields surrounding Earth and Jupiter.

Another component of Saturn's magnetic field is an extensive ring of current of 10 million amperes flowing from west to east. Near the axis, the ring-current field is not parallel to the main field but becomes parallel with distance. Moreover, field lines are outward bound above the equatorial plane and inward bound below that plane. Thus, Saturn's magnetic field is made up of the main dipole field and the ring-current field, along with the boundary currents of the magnetopause (the border at which the solar wind meets the planetary magnetic field), thereby defining the magnetosphere. Because of the superposition of the ring-current field on the planet's intrinsic dipole field, the shape of



An artist's diagram of Saturn's magnetic field. (ESA) the overall magnetosphere is stretched outward along the equatorial plane, creating a bulge that distorts the poloidal aspect of the field.

As confirmed by Pioneer 11, the Voyagers, and the Cassini orbiter, interaction between Saturn's magnetic field and the solar wind causes a magnetosphere accompanied by a well-defined bow shock, a plasma sheath, and a magnetotail, similar in many respects to those associated with Earth and Jupiter. The shape of the magnetosphere can be schematically represented by a paraboloid of revolution around the Saturn-Sun line. The size of the magnetosphere depends on the varying pressure of the solar wind. If this pressure is low, the magnetopause occurs at a greater distance, thus inflating the magnetosphere. The solar activity cycle, interplanetary conditions, and local variations influence solar wind pressures. Magnetospheric and myriad complex radiative processes that occur within it result primarily from the interaction between charged particles of the solar wind and Saturn's intrinsic magnetic field. It is assumed that the transfer of energy from the solar wind to the magnetosphere occurs within the outer layer of the magnetosphere and the magnetotail. The magnetosphere effectively entraps, stores, and reradiates the energy of the solar wind.

With regard to Saturn's magnetic field, Pioneer 11 and the Voyager probes found that there was a near alignment of the magnetic and rotational axes and a displacement of the magnetic dipole center compared to the planetary center. They also discovered the existence of an additional field produced by a ring of current flowing west to east in the equatorial plane. Because of the varied areas covered by these probes' differing trajectories and the complementary data they provided, the Pioneer and Voyager missions provided a fairly accurate model of the Saturnian magnetic field configuration. Cassini's magnetometer filled in a great many gaps and found some surprises, especially with regard to the satellite Enceladus.

Saturn's ring systems are intricate, elaborate, and fascinating both in appearance and from a scientific point of view. The rings are thin and known to be composed of pieces of ice and chunks of rocky material, most of which are probably much smaller than ten meters in diameter. Stars can dimly be seen through the rings. The entire system rotates in the equatorial plane, within about 180,000 kilometers of the center of the planet. All of Saturn's major satellites, starting with Mimas, are located beyond the principal ring systems. Smaller ones are embedded and inside the ring systems.

The regions surrounding the major rings of Saturn are devoid of trapped charged particles. Up to a radial distance of about 150,000 kilometers, the density of charged particles is found to be negligible. Thereafter it rises abruptly, with intermittent variations caused by the absorption of particles by the satellites and the tenuous E ring. The maximum density of ions and electrons is observed between the orbits of Tethys and Rhea, between 300,000 and 60,000 kilometers from Saturn's center. The same region, which is rich in dense plasma, is lacking in low-energy electrons and protons. They are most likely absorbed by the neighboring satellites.

In this inner magnetosphere, spinning rapidly, charged particles are coupled to the magnetic field, resulting in a 60,330-kilometer-thick plasma sheet about 240,000 to 422,000 kilometers from the planet. From this distance to the orbit of Titan, there is a vast torus of neutral hydrogen, produced by the photochemical breakdown of methane escaping from the giant satellite's atmosphere. Another source of charged particles is provided by the collision of neutral hydrogen atoms with the magnetosphere of Titan.

Saturn's magnetospheric radio emission has three components, all of which are presumed to originate in the magnetic plasma surrounding the planet. Its kilometric radiation was found to have a fundamental periodicity of 10 hours and 40 minutes, presumably the rotational period of the magnetosphere. These magnetic storms seem to affect and are correlated with solar wind conditions, as well as the relative position of the satellite Dione. Saturn's kilometric radio emissions, first discovered by Voyager 1 in January, 1980, exhibit a wide range of period modulation (10.66 hours, 66 hours, and 22 days) along with intensity and power variations. The axial symmetry of Saturn's magnetic field does not yield the kind of intensity modulation that is associated with the wobbly motion of Jupiter's field. Saturn appears to radiate greater power toward its dayside than toward the nightside; this phenomenon has its origin in midlatitude to polar cusps, regions where magnetic field lines enter or leave the planet. Magnetospheric kilometric radiations—short bursts of radio emissions caused by electrostatic discharges ranging in frequency from 20 kilohertz (kHz) to 40 megahertz (MHz) with a periodicity of 10 hours and 10 minutes—were also detected by the planetary probes. The total power radiated by these discharges appears to be comparable to those of the kilometric emissions. The plasma trapped within the magnetosphere also has radio emissions, with a frequency exceeding 2 to 3 kHz and a recurrence period of 10 hours and 40 minutes. These low-frequency radio waves were detected by both Voyager probes. The total power radiated is estimated to be on the order of a million watts.

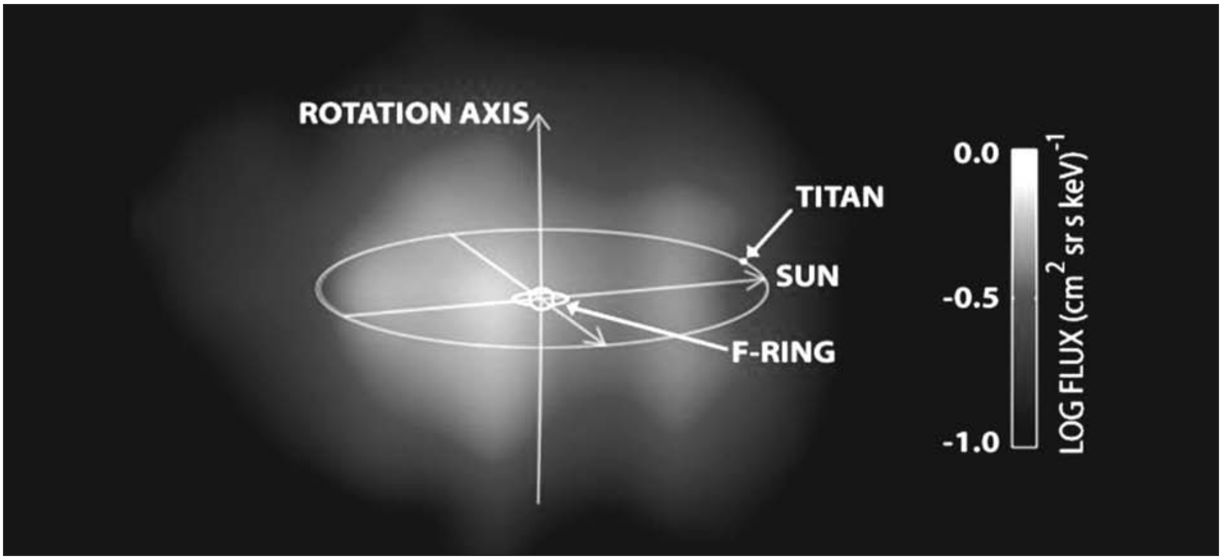
A dense atmosphere combined with a strong magnetic field indicates the presence of auroras. Voyager l's detection of strong ultraviolet radiation above 76° north latitude and below 78° south latitude confirms the existence of Saturnian auroras similar to those in the polar regions of Earth and Jupiter, with possible periodic longitudinal variation in intensity. Based on these measurements, the power required to generate the Saturnian auroras is estimated at 200 billion watts, which is seven times the corresponding value for Earth's auroras. The main source of this energy is the interaction between the solar wind and Saturn's magnetosphere. Confinement of charged particles within the magnetic field and their being forced to travel along intense magnetic field lines near the polar regions cause the auroras. More intense study of Saturn's auroral activity was effected using the Hubble Space Telescope and from the orbiting Cassini spacecraft.

Because Saturn's magnetic field is highly symmetrical, there is a well-defined bow shock and a steady magnetopause. Saturn has the necessary ingredients for the presence of an active outer magnetic field: a planetary dipole field, a ring-current field, and contributions to the field from a magnetopause and tail currents. The outer magnetosphere of Saturn is supplied with a constant flow of hydrogen and nitrogen plasma by Titan and its magnetic field. This plasma torus is co-rotating with the magnetosphere, undergoing convective motion at the same time, and is estimated to have a temperature of 1 million kelvins—the hottest of planetary surroundings known to date. Neutral atoms escaping from Titan are photoionized, becoming part of the plasma torus. It is surmised that the co-rotations of plasma, combined with the continual radial movement of plasma torus caused by fluctuations in the pressure of the solar wind, are responsible for the plasma's unusual heating, which leads to an increase in high-energy particles. In addition to the elevated plasma temperature that results in high-energy particles, there is a coincident rise in field strength. This jump in field strength appears to be sharper when the magnetosphere contracts. Voyager data indicate discontinuities between the region inside the magnetosphere and the outside, where the solar wind persists. An acceleration of electrons and ions occurs in the magnetotail. Voyager 1 detected low-energy ions streaming toward, and high-energy ions racing away from, Saturn at distances of 2 million to 2.7 million kilometers.

The bow shock, the magnetopause, the acceleration of ions in the magnetotail, a spectrum of energy, and the flux of charged particles appear to be common to the outer magnetospheres of Saturn and Earth. A major difference is the Titan-fed hot plasma torus of Saturn's outer magnetosphere. There are also many unexplained phenomena in Saturn's magnetosphere, one of them being the observed fluctuation in temperature at altitudes between 600,000 and 900,000 kilometers. It is not known whether the magnetotail drains the angular momentum of Saturn through ejection of high-energy particles.

KNOWLEDGE GAINED

The search for Saturn's magnetic field began in 1955, with the accidental discovery of decimetric radio emissions from Jupiter. It was surmised that radio emissions from Saturn's magnetic field would be weaker than those of Jupiter and hence not observable by ground-



This Cassini image from June 21, 2004, shows Saturn's magnetosphere: The Magnetospheric Imaging Instrument, by detecting the hydrogen atoms that escape the magnetosphere, revealed a glow that reflects the area of magnetospheric influence around the planet. (NASA/JPL/Johns Hopkins University)

based radio telescopes. Long-wave radio bursts coming from the direction of Saturn were detected in 1975 by the Interplanetary Monitoring Platform (IMP) 6. Instruments on board Pioneer 11 in 1979, however, confirmed the existence of an extensive magnetic field surrounding Saturn. Detailed observations were carried out by more sophisticated instrumentations of Voyagers 1 and 2. Long-term studies of the magnetic fields and particle environment around Saturn continue to be performed by the Cassini spacecraft.

Data from Pioneer 11 and the two Voyagers suggest that the magnetic field of Saturn originates at a greater depth than does Jupiter's. Unlike the fields of Earth and Jupiter, Saturn's magnetic field is highly axisymmetric because of the overlapping of the dipole and the rotational axes. The main field is hemispherically asymmetrical because of a northward offset of the dipole axis. Field strength at an altitude of 60,330 kilometers is 0.2 gauss, roughly a third of the equatorial geomagnetic field. A second component of the magnetic field of Saturn is the extensive current ring that flows from west to east. The boundary currents of the magnetopause constitute a third component of the field.

The inner magnetosphere of Saturn has been found to be free of charged particles. The maximum density of ions and electrons occurs between the orbits of Tethys and Rhea. Between the orbits of Tethys and Titan, there is a vast torus of neutral hydrogen produced by the photochemical breakdown of methane escaping from Titan's atmosphere. Saturn's kilometric radio emission, along with a variety of periodic and nonperiodic bursts of radiation, is presumed to originate in the magnetic plasma surrounding the planet. The planet's kilometric radiation has a periodicity of 10 hours and 40 minutes. Short bursts of radio emissions have been traced to electrostatic discharges.

Pioneer 11 and the two Voyagers also provided tantalizing information about Saturn's auroras, phenomena that occur invariably in the presence of a dense atmosphere combined with a strong magnetic field. Saturn's auroras, which occur above 76° north latitude and below 78° south latitude, are powered by the interaction between the solar wind and the magnetosphere.

Cassini's magnetometer naturally was capable of determining the direction and strength of Saturn's magnetic field with the best technology available at the time the spacecraft was designed. Called a dual technique magnetometer, it also was able to assist in the determination of the size and nature of Saturn's core. In order to have the sensitivity necessary for its intended research, the magnetometer assembly included a flux gate magnetometer and a vector/scalar helium magnetometer placed along an 11-meter-long boom to avoid interference from spacecraft electronics. The ability to determine three-dimensional magnetic field maps would be used not only on Saturn but also to search for magnetic fields of the ringed planet's satellites. The magnetometer was intended to make new discoveries and to answer some of the outstanding questions raised by the Voyager results.

Cassini confirmed the Pioneer 11 and Voyager magnetic field findings, also returning some surprising insights about Saturn's satellite Enceladus. Much remains to be learned about this complex miniature-solar system during Cassini's extended mission, which was authorized in 2008, given the orbiter's good health, after Cassini's primary mission ended.

CONTEXT

Study of Saturn's magnetic field began with radio astronomers, whose findings and conjectures were to play a dominant part in the planning of planetary exploration probes of the 1970's. Pioneer and Voyager data vastly advanced scientists' understanding of planetary magnetic fields in general and Saturn's field in particular. For example, the axisymmetric nature of Saturn's magnetic field and its effects on magnetospheric processes make it quite different from both Earth's and Jupiter's fields. However, data show that Saturn's magnetosphere (like those of Earth, Jupiter, and Venus) has an extended magnetotail, in which interaction between solar plasma and the magnetic field produces a host of phenomena yet to be researched. The study of magnetospheric processes in general provides a basis for understanding cosmic plasma.

Saturn's magnetic field is basically dynamo-driven, like similar fields in the solar system, including even that of the Sun. Close to the planet, plasma depletion is caused by the ring system closer to the planet, but there is a vast torus of Titan-fed hot plasma at the outer magnetosphere. Saturn appears to possess three distinct regions of plasma: an inner plasma torus, an extended plasma sheet, and the hot outer plasma torus. Both the temperature and the thickness of the plasma disk increase with distance from the planet. The nature of the sources and sinks of these plasma regions is yet to be determined, although speculations abound. The plasma torus is co-rotating, with velocities decreasing by 10 to 20 percent beyond 480,000 kilometers of altitude.

Based on knowledge of Earth's magnetic field and the data provided by Pioneer 11, Voyagers 1 and 2, and Cassini, a model of Saturn's field has been constructed. There are undoubtedly numerous processes occurring in its distant magnetosphere that need to be further examined. The variation in the size and shape of the magnetosphere needs to be examined over an extended period. Quantitative studies of the plasma flow from the satellites of Saturn (Titan and Enceladus in particular) and research into the charge absorption properties of the ring system are projects of long-term interest as well.

The theoretical model of Saturn's magnetic field is based on a solid foundation provided by spacecraft data. Yet there are numerous unanswered questions and doubts, and the model must be further refined. Continuing research into the atmospheres and magnetic fields of planets such as Jupiter and Saturn is vital. However, no follow-on mission to Saturn beyond Cassini was advanced and approved by the time that Cassini's flight was extended two additional years after completion in 2008 of the orbiter's primary mission objectives.

V. L. Madhyastha

FURTHER READING

Alexander, Arthur Francis O'Donel. *The Planet Saturn: A History of Observation, Theory, and Discovery*. New York: Macmillan, 1962. An excellent historical reference for all readers, this book details the study of Saturn up to the mid-twentieth century. Organization is somewhat choppy, however, rendering some sections difficult to read.

Bortolotti, Dan. *Exploring Saturn*. New York: Firefly Books, 2003. A look at the Cassini-Huygens mission for a younger audience. Full of charts, photographs, a section on observing Saturn, and a history of our understanding of the Saturn system from antiquity to the launch of Cassini.

Gehrels, Tom, and Mildred Shapley Matthews, eds. *Saturn*. Tucson: University of Arizona Press, 1984. Many of the articles collected here are by individuals who were involved with the Pioneer 11 or Voyager 1 and 2 projects; some of them were also responsible for publishing the first reports of the projects' scientific findings. Thus, the work is authoritative, and though it is intended for the specialist, the general reader can obtain from it much useful information, including helpful lists of references.

Harland, David M. *Cassini at Saturn: Huygens Results*. New York: Springer, 2007. The cover illustrates a landing site on Titan. The text inside provides a thorough explanation of the entire Cassini program, including the Huygens landing on Saturn's largest satellite. Essentially a complete collection of National Aeronautics and Space Administration's releases from the start of Cassini flight operations through the majority of Cassini's seventy orbits during its primary mission, which concluded a year after this book was published. Technical but accessible to a wide audience.

_______. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002. A technical description of the Cassini program, its science goals, and the instruments used to accomplish those goals. Written before Cassini arrived at Saturn, it nevertheless provides a historical review of pre-Cassini knowledge of the Saturn system.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. Covers the Pioneer 11 and Voyager data concerning Saturn and previews Cassini mission objectives. Takes a comparative planetology approach rather than approaching each moon and planet in separate chapters.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science, this volume focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs, the work is accessible to the serious general audience.

Morrison, David. *Voyages to Saturn*. NASA SP-451. Washington, D.C.: Government Printing Office, 1982. This is a well-written account of the encounters of Voyagers 1 and 2 with Saturn and a summary of the resulting scientific findings. Containing a wealth of information that is probably not available to the general reader elsewhere, this volume is a primary resource for Voyager returns on Saturn.

Russell, Christopher T. *The Cassini-Huygens Mission: Orbiter Remote Sensing Investigations*. New York: Springer, 2006. Provides a thorough explanation of the remote-sensing investigations of the orbiter and lander. Outlines the scientific objectives of all instruments on the spacecraft and describes the planned forty-four encounters with Titan. Given the publication date, only early science returns are covered.

Schardt, A. W. "Magnetosphere of Saturn." *Review of Geophysics and Space Physics* 21 (1983): 390-402. This historic review article briefly describes Saturn's magnetic field and its associated phenomena in the light of findings from Pioneer 11, Voyager 1, and Voyager 2. The author presents a brief history of the projects and assumes minimal technical knowledge on the part of the reader. The article includes an extensive list of references.

See also: Auroras; Brown Dwarfs; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Jovian Planets; Jupiter's Interior; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits: Couplings and Resonances; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Saturn's Atmosphere; Saturn's Interior; Saturn's Ring System; Saturn's Satellites.

Saturn's Ring System

Categories: Planets and Planetology; The Saturnian System

Data transmitted by Pioneer 11, and in greater detail by Voyagers 1 and 2, revolutionized the understanding of Saturn's complex ring system previously obtained by observations from Earth-based telescopes This information revised models based on those earthbound investigations dating back to the discovery of the rings more than three centuries ago. Hubble Space Telescope and Cassini orbiter studies then built upon data returned by earlier spacecraft and revealed an even more complex ring system at Saturn.

OVERVIEW

Pioneer 11, Voyager 1, and Voyager 2—three deep space probes launched between 1973 and 1977—all encountered Jupiter before their trajectories were directed toward Saturn. Pioneer 11 was also known as Pioneer Saturn during its Saturn flyby, as it was a pathfinder for the more sophisticated two Voyager spacecraft. Voyager 2, launched after Pioneer, would have to pass through Saturn's ring plane at a distance of 2.86 Saturn radii (about 112,000 kilometers above the surface) in order to be put on a trajectory for a potential Uranus flyby. Although beyond the main rings, there was the possibility that a tenuous ring existed in this region. Such a ring would pose a threat to any spacecraft passing through at high speed. The decision was made to have Pioneer 11 cross the ring plane at this distance to determine whether it was safe for the more valuable Voyager spacecraft to come.

The Cassini spacecraft was launched on October 15, 1997, and after a series of gravity assists arrived in orbit about Saturn on July 1, 2004. The primary mission of this spacecraft was completed in four years, but fortunately funding was available for an extended examination of the Saturn system as the spacecraft remained in near-perfect health at the end of its primary mission.

Pioneer 11 survived the crossing of the ring plane on September 1, 1979, at a distance of 2.82 Saturn radii. The success of this maneuver was evident from the continued reception of the spacecraft's radio transmissions before, during, and after the ring crossing. Pioneer 11 then swung around the planet, crossing the ring plane a second time (about 2.5 hours after the first crossing) at a distance of 2.78 Saturn radii. Again, although it took some hits, there was no detectable damage during this crossing.

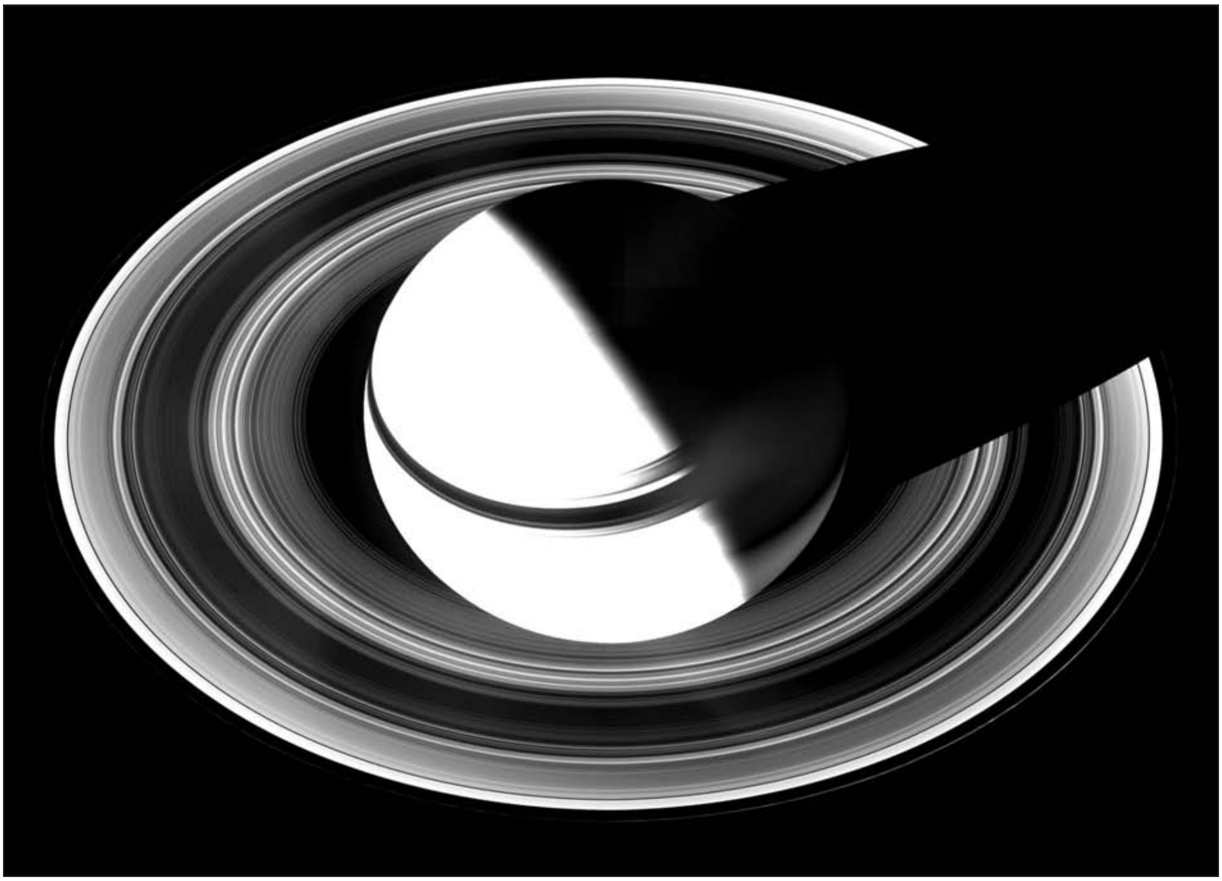
During its Saturn flyby, Pioneer 11 transmitted images of the rings made by an imaging photopolarimeter, an instrument which produces images by means of polarized light. In this case it made images at two visible wavelengths, one in the red region of the electromagnetic spectrum and another in the blue. These images could be processed at the receiving station to simulate color pictures. As the spacecraft approached Saturn, the resolution exceeded that of earthbound observations. At a distance of about 1 million kilometers, a new ring was detected in one of the images. It was named the F ring. It was too narrow and too close to the outer edge of the A ring to have been seen from Earth. Pioneer 11 also collected data on Saturn's rings by using transmitted sunlight, a method not possible for observations made from Earth. With this method, images are analogous to a photographic negative; gaps in the rings appear bright instead of dark, while the dense, bright parts of the rings appear dark. A week after the first encounter, the Saturn flyby was essentially complete, and Pioneer 11 was on its way on an escape trajectory that would take it out of the solar system.

As the planet moves about the Sun, the orientation of the ring plane relative to Earth varies. In early 1980, prior to the Voyager 1 flyby, Saturn's rings appeared sideways in Earth-based observations. The main rings were practically invisible, because they appeared so thin from that perspective. Conditions were favorable, however, for detecting faint satellites and diffuse rings which would otherwise be lost in the glare of the main rings. A faint ring was detected beyond the F ring by astronomers during this period. This ring, now called the E ring, had actually been discovered in 1966, the last time that the rings were edgewise, by a then-unknown astronomer, Walter A. Feibelman. At that time, its existence was called into question by other observers. Pioneer 11 vindicated Feibelman.

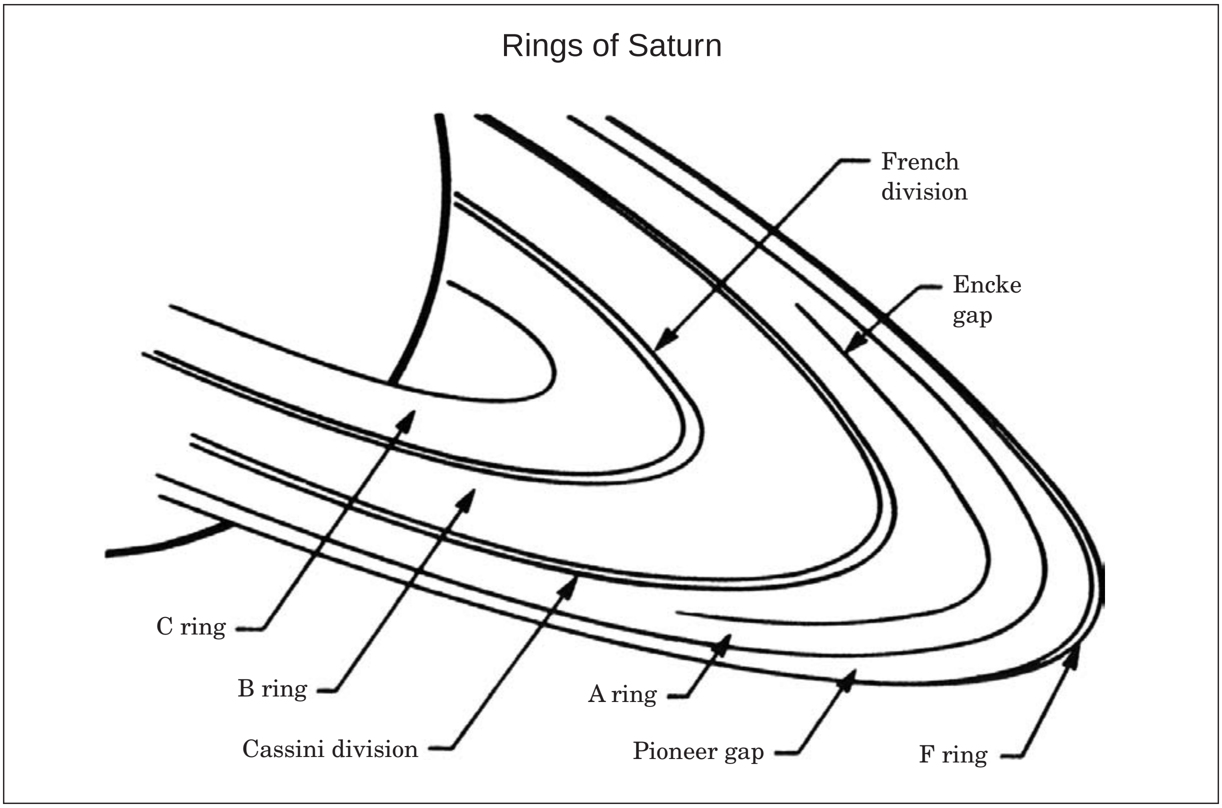
The Voyager imaging system represented a tremendous improvement over that of Pioneer 11. It consisted of two television cameras. One was outfitted with a wide-angle lens while the other had a narrow-angle lens. These cameras were mounted on a scan platform which could be aimed continuously at the target, a process referred to as slewing or motion compensation. This prevented smearing of images due to the rapid motion of the Voyager spacecraft, and was accomplished without using propellant to reorient thrusters to move the spacecraft. Pioneer 11's imaging photopolarimeter had to rotate with the spin-stabilized spacecraft, and could only record the subject once during each revolution. Voyager images could be transmitted more rapidly in black and white or in color. They also had higher resolution.

During October, 1980, Voyager 1 was traveling toward Saturn at an average rate of 1.3 million kilometers per day. At the end of the month it was 17 million kilometers from the planet, and the improved resolution brought to light new details within the rings. The principal rings, particularly the B ring and C ring, were found to consist of narrow concentric rings. The ringlets, as they were called, proved to be so numerous that they suggested the analogy of grooves on a phonograph record. The main gap in the ring system, the Cassini division between the A ring and B ring, was not a total void. It contained some ringlets of its own.

The B ring exhibited some curious radial



Saturn and its rings, compiled from thirty-six images taken as the Cassini spacecraft approached on January 19, 2007, about 1.23 million kilometers from the planet. The planet casts a shadow on the rings facing away from the Sun. (NASA/JPL/Space Science Institute)



Source: David Morrison, *Voyages to Saturn*. NASA SP-451. Washington, D.C.: National Aeronautics and Space Administration, 1982, p. 25.

streaks referred to as spokes which rotated with the ring while changing their shapes. This was a completely baffling phenomenon. In order to study them more carefully, Voyager 1 was programmed to make images every five minutes over a period of ten hours. In the process, two small satellites were discovered. One orbited along the outer edge of the F ring, the ring that had been discovered by Pioneer 11, while the other orbited along the ring's inner edge. The specific location of these satellites is considered to be essential to the permanence of the F ring. The F ring itself also has a curious property. It is braided.

During its encounter with Saturn's rings on November 12, 1980, Voyager 1 swung around Saturn and crossed the ring plane twice without damage. Before being directed to escape the solar system, it transmitted spectacular images of Saturn and its rings. The number of ringlets that could be detected with the Voyager instrumentation was estimated to have an upper limit of one thousand. Voyager 1 discovered two more rings, the D ring, which replaced the C ring as the innermost ring, and the G ring, beyond the F ring discovered by Pioneer 11. Evidence for the G ring had been obtained by Pioneer 11. Its charged particle detector registered a decrease in intensity which could have been caused by absorption by the ring particles. Voyager 1 confirmed the existence of the E ring by photographing it directly.

As a result of discoveries made by Voyager 1, the program of Voyager 2 was revised for focused studies of the rings. A decision was made to program the second Voyager to collect data during a stellar occultation of Saturn's rings. During the occultation, Voyager 2 would be positioned above the ring plane, focusing on the selected star through Saturn's rings. Delta Scorpii, the star, was bright enough to be seen through the partially transparent rings. One of

Voyager's instruments, the photopolarimeter subsystem (PPS), would be focused on the star for about two hours. The PPS would not form images. It would measure the rapid fluctuations in the brightness of the star. These fluctuations would be caused by the ring particles in the line of sight momentarily blocking the light from the star. Data could be analyzed to map variations in ring structure. Voyager 2 also scanned the area for small satellites which might be embedded within the rings.

As Voyager 2 approached Saturn, images of the rings achieved an extraordinary resolution of ten kilometers. Still, there was no evidence of embedded satellites. The number of observed ringlets increased with the improved resolution. On August 25, 1981, the occultation of Delta Scorpii was successfully observed. As seen from Earth, Voyager then proceeded to cross the ring plane behind Saturn. The successful crossing could not be confirmed until radio communications were resumed about one hour later. Voyager 2 was then directed toward Uranus. Shortly afterward, ground controllers discovered that the scan platform had lost its azimuth motion and had been shut down by the on-board computer. Many of the programmed images had not been acquired, although some good images of the F ring were obtained before the scan platform failed. Quick actions by the Voyager team salvaged much of the anticipated scientific investigations as Voyager 2 left Saturn.

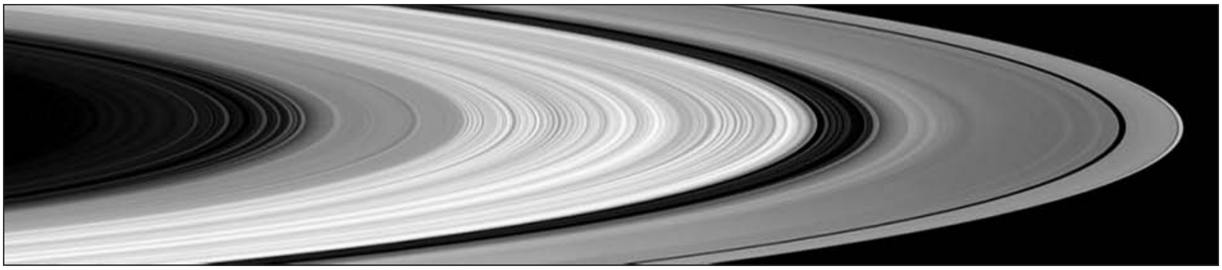
As Cassini approached orbit about Saturn, it made its closest planned encounter with the planet's rings. Looking at the B ring Cassini noted a puzzling lack of spokes presently. This strongly suggested that the spoke phenomenon might be a seasonal effect. Later in the mission Cassini produced numerous images of the spokes.

Cassini also examined gaps in the rings and noted far more structure in them than had the Voyagers, and found additional shepherding satellites. For example, inside the 42-kilometer-wide Keeler gap within the A ring, in May, 2005, Cassini discovered a small satellite. This body clears out material within this gap. In 2006 Cassini imaged a very faint dust ring located near the satellites Janus and Epimetheus. It is believed that this 5000-kilometer-wide ring is composed of particles liberated by meteoroids

Rings of Saturn

	Radius (km)	*Radius/ Eq. Radius*	*Optical Depth*	*Albedo ($\times 10^{-3}$)*	*Surface Density (g/cm^2)*	*Eccentricity*
Saturn equator	60,268	1.000	—	—	—	—
D ring	>66,900	>1.11	—	—	—	—
C inner edge	74,658	1.239	0.05-0.35	0.12-0.30	0.4-5	—
Titan ringlet	77,871	1.292	—	—	17	0.00026
Maxwell ringlet	87,491	1.452	—	—	17	0.00034
B inner edge	91,975	1.526	0.4-2.5	0.4-0.6	20-100	—
B outer edge	117,507	1.950	—	—	—	—
Cassini division	—	—	0.05-0.15	0.2-0.4	5-20	—
A inner edge	122,340	2.030	0.4-1.0	0.4-0.6	30-40	—
Encke gap	133,589	2.216	—	—	—	—
A outer edge	136,775	2.269	—	—	—	—
F ring center	140,374	2.329	0.1	0.6	0.0026	—
G ring center	170,000	2.82	1.0×10^{-6}	—	—	—
E inner edge	~180,000	3	1.5×10^{-5}	—	—	
E outer edge	~480,000	8	—	—	—	—

Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.



This mosaic of six images from Cassini (taken December 12, 2004) displays gaps, gravitational resonances, and wave patterns across about 62,000 kilometers of the ring plane. (NASA/JPL/Space Science Institute)

impacting these two satellites. That same year another faint dust ring was found, this one existing close to the satellite Pallene. This one is only 2500 kilometers in radial extent. It too is suspected to be composed of particles generated by collisions of objects, specifically with Pallene.

KNOWLEDGE GAINED

Three of Saturn's major rings (D ring, F ring, and G ring) were discovered with the help of deep space probes. The existence of the faint E ring was confirmed. The rings, in order of increasing distance from the planet, are D, C, B, A, F, G, and E. Saturn's rings were first named in alphabetical order in 1850: A, B, and C for the three rings then known. The convention now is to name the rings in the order of their discovery.

Three narrow gaps in the rings were also found: one in the C ring, one at the inner edge of the Cassini division, which separates the A ring and B ring, and a third one in the A ring, close to its outer edge.

While individual ring particles have not been observed directly, the distribution of particle sizes can be estimated from measurements of the transmission of light and radio waves through the rings and scattering, or reflection, from the rings. Being an orbital platform, Cassini was able to conduct multiple radio occultation measurements. The sizes range from microns to about ten meters. Spectroscopic measurements confirm that these particles are primarily composed of ice, with some impurities. Considering the larger-sized particles only, scientists estimate that the thickness of the rings to be about one hundred meters. This is relatively thin since the ring system is thousands of kilometers wide.

Rings discovered by deep space probes differ from the ones first identified by ground-based observation. The outer two rings, G and E, are composed of particles in the micron range and do not have the ringlet structure. They are diffuse and are probably thicker than one hundred meters. It is believed that Voyager 2 passed through the G ring when it was leaving Saturn. Its plasma wave detector recorded a large increase in intensity in the vicinity of the ring. The increase was apparently caused by the spacecraft colliding with ring particles as it traveled through the ring at 10 kilometers per second. Impacts would have vaporized the particles, producing puffs of ionized gas (plasma) which would have been recorded by the detector.

Braided strands of the F ring observed by Voyager 1 were found by Voyager 2 to have changed to parallel strands. They changed back to braided strands by the end of the encounter, one example of the dynamic nature of the rings. Two so-called shepherd satellites serve to confine the ring. They are elongated rather than spherical. The long axis of the larger one is about 140 kilometers. Cassini was able to observe the dynamic behavior of the F ring over the course of its many orbits about Saturn.

Of all the ring components seen in the outer solar system, Saturn's F ring displays the most unusual and dynamic activity. Features in it can be seen to change on timescales ranging from just hours to several years. Studies using Cassini's Ultraviolet Imaging Spectrometer during stellar occultations, found at least a dozen objects within the F ring ranging from 27 to 10,000 meters in size. Data suggested that these features were aggregates that had temporarily clumped together, the supposition being that within the rings material coalesces and breaks apart due to gravitational and collision processes, respectively. In mid-2008 researchers using Cassini observations, published a paper in *Nature* that attributed the unusual characteristics of F ring dynamics to perturbations created by small moonlets, making the F ring the ring location in the solar system found thus far where a serious number of collisions happen regularly. In some ways F ring dynamics provide a window into the early solar system's protoplanetary disk at a time when collisions of small particles were needed to drive planetary formation.

In addition to the gravitational influence of the satellite Prometheus, a moonlet on the order of only a few kilometers in radius appears to be colliding with F ring particles orbiting close to the core of the ring generating fan-shaped structures. Another combination of gravitational effects and collisions produces structures referred to as jets. A Cassini image taken in 2004 appears to have identified a five-kilometer-wide moonlet that may produce some of the largest of the observed jets in the F ring. These phenomena would continue to be studied throughout the extended Cassini mission.

The D ring appeared to be a collection of ringlets too faint to be seen from Earth. It has a relatively small percentage of micron-sized particles. Analysis of Voyager data revealed three ringlets. These were designated D68, D72, and D73 in increasing distance radially from the planet toward the C ring. Here again, Cassini revealed dynamic behavior and more complex structure. Twenty five years after its discovery, D72 was observed to be fainter and at a location 200 kilometers closer to Saturn. The gap between D73 and the C ring was not empty. Cassini recorded fine structure with ripples of material 30 kilometer apart.

The Cassini division, when viewed in transmitted light, appeared to have five rings in its central region and a gap on each side separating it from the A ring and B ring. The gap at the inner edge has an eccentric ringlet, which was scanned for small satellites that might confine it. None was found, but Cassini uncovered even more structure in the gap than revealed by the Voyagers.

The B ring, the largest and brightest of the "classical" rings, has the most elaborate structure, literally thousands of ringlets. Spokes are transient features that appear bright in transmitted light but dark when viewed from the sunlit side. They consist of micron-sized particles. When electrically charged, they interact with Saturn's magnetic field, a process that explains some of their properties. There is evidence that the ringlets themselves are manifestations of a wave, propagating through the ring in the form of a spiral. This spiral might, in turn, produce the observed density variations. Cassini provided a great deal of evidence for not only density waves in the rings but also torsion waves.

CONTEXT

History records that in 1610 Galileo was the first person to observe Saturn through a telescope. He described it as having a close satellite on either side. Later observers used the Latin word "ansae" (handles, in the sense of cup handles) to describe what they saw. This term is still used to refer to the parts of the rings that are visible on either side of the planet. Christiaan Huygens, based on observations in 1655, concluded that Saturn was surrounded by a thin ring not attached to the planet. Gian Domenico Cassini, discovered the main gap in the ring system in 1675, showing that there were two rings, now called A and B. The C ring was discovered in 1850.

In 1867, Daniel Kirkwood, an American astronomer, applied a resonance theory he had developed to explain the existence of the Cassini division. A ring particle in the division orbits Saturn with a period one-half that of Saturn's satellite Mimas, which is farther from the planet. Every other period, the particle passes Mimas in the same part of its orbit and is affected by a gravitational force pulling it outward. This periodic or resonant force would clear the ring of particles. The process is actually more complicated. Additional satellites have to be considered, and the gravitational interaction is not as simple as described. In some instances, the force might produce a ringlet by causing a spiral density wave to form, as mentioned with respect to the B ring. Originally, this type of wave was introduced to explain the structure of spiral galaxies, such as the Milky Way. Just as the complexity of the ring system was completely unexpected, the concept that Saturn's rings may have some similarities to spiral galaxies, despite their enormous difference in size, could not have been anticipated before the Saturn encounters.

Prior to Pioneer 11, Voyagers 1 and 2, and Cassini, Saturn's rings could be described only in relatively simple terms. It had been established that the rings consist of discrete particles in orbit around the planet. Spectroscopic measurements showed that the inner parts of the rings rotate faster than the outer parts. Thus, it was clear that the rings were not solid disks.

In addition, it had been proved theoretically that any natural satellite large enough to be held together simply by the force of its own gravity would be fragmented by tidal forces exerted on it by the planet if it was closer than about 2.4 times the radius of the planet. This inner limit is known as the Roche limit, named for the nineteenth century French mathematician Edward Roche. All rings, except the F ring and E ring, are at a distance greater than 2.4 Saturn radii away from the planet. Roche suggested that the rings were formed by fragmentation of a satellite which came too close to the planet. Spacecraft data support another possibility. A number of Saturn's icy satellites were found to be pockmarked by impact craters. Ring particles could be remnants of the debris resulting from the collisions that produced such craters.

Voyager 2 successfully completed its Uranus encounter on January 24, 1986, and continued on its route to Neptune passing through that system in August, 1989. This spacecraft is unique in having made observations at close range of the four known planetary ring systems: those of Jupiter, Saturn, Uranus, and Neptune. The extensive data set from the Saturn encounters made by four spacecraft form the basis for a unified model of planetary rings in general.

With the extension of the Cassini mission in 2008, there came the promise of new and exciting discoveries about the complex ring system of Saturn for several more years to come. One such unexpected investigation would be follow-up studies of a possible ring system about Saturn's large satellite Rhea. Cassini in November, 2005, was directed to use its Magnetospheric Imaging Instrument (MIMI) to observe the planet's magnetosphere in the vicinity of Rhea. MIMI found three specific diminishments of energetic particles located symmetrically about either side of this satellite, Saturn's second largest. The supposition was that there might be three rings made of particles on the order of a meter in size existing with an equatorial dish of material. This was not the only possible explanation, but if Cassini data confirmed it to be correct, then Rhea would be the only satellite known to have its own ring system.

Howard L. Poss

FURTHER READING

Alexander, A. F. O. *The Planet Saturn: A History of Observation, Theory, and Discovery*. New York: Dover Publications, 1980. The standard work on the history of observations of Saturn, from ancient times to 1960. A large part of the book is devoted to the rings. The author frequently uses quotations from the original sources. A useful reference. Illustrated with drawings and photographs. Contains an extensive index.

Consolmagno, Guy. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994. A text accessible to college-level science and nonscience readers alike. Presents subjects at low-level mathematics and also involves integral calculus where required. Demonstrates how the area of planetary science progresses by questioning previous understanding in light of new observations.

Elliot, James, and Richard Kerr. *Rings: Discoveries from Galileo to Voyager*. Cambridge, Mass.: MIT Press, 1984. The discovery of Uranus's rings is described. Other topics include the discovery of Jupiter's ring and the Saturn encounters. For college and high school students with background in the physical sciences. Includes an extensive bibliography of journal articles. Illustrated.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that have investigated all planets with the exception of Pluto by the publication date. Takes an astrophysical approach to give our solar system a wider context as just one member of similar systems throughout the universe.

Harland, David M. *Cassini at Saturn: Huygens Results*. New York: Springer, 2007. The cover illustrates a landing site on Titan. The text inside provides a thorough explanation of the entire Cassini program including the Huygens landing on Saturn's largest satellite. Essentially a complete collection of National Aeronautics and Space Administration (NASA) releases from the start of Cassini flight operations through the majority of Cassini's seventy orbits of its primary mission. Cassini's primary mission concluded a year after this book entered print. Technical writing style but accessible to a wide audience.

_______. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002. Another book in Springer's Space Exploration Series, this is a technical description of the Cassini program, its science goals and the instruments used to accomplish those goals. Written before Cassini arrived at Saturn. Provides a historical review of pre-Cassini knowledge of the Saturn system.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Available to the serious general audience.

Lovett, Laura, Joan Harvath, and Jeff Cuzzi. *Saturn: A New View*. New York: Harry N. Abrams, 2006. A coffee-table book replete with about 150 of the best images returned by the Cassini mission to Saturn. Includes the planet, its many satellites, and the complex ring systems.

Morrison, David. *Voyages to Saturn*. NASA SP-451. Washington, D.C.: Government Printing Office, 1982. The official account of the Saturn missions. The author is an astronomer and was a member of the Voyager Imaging Science Team. An introductory section describes the revival of interest in planetary astronomy, stimulated by the advancement of space technology. Descriptions of the Pioneer 11 and Voyager Saturn encounters are provided. The book ends with a section on Saturn's rings. Well illustrated. Contains a glossary of terms.

Russell, Christopher T. *The Cassini-Huygens Mission: Orbiter Remote Sensing Investigations*. New York: Springer, 2006. Provides a thorough explanation of the remote sensing investigations of the Orbiter and lander. Outlines the scientific objectives of all instruments on the spacecraft. Describes the planned 44 encounters with Titan. Only provides early science return.

Washburn, Mark. *Distant Encounters: The Exploration of Jupiter and Saturn*. New York: Harcourt Brace Jovanovich, 1982. Similar to the book by Cooper but more extensive in scope. Washburn describes the atmosphere among the researchers involved with the Voyager program. Astronomical data are provided, and the budgetary problems that the program faced are discussed. Illustrated with quality photographs.

See also: Auroras; Brown Dwarfs; Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Extraterrestrial Life in the Solar System; Jovian Planets; Jupiter's Interior; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Orbits: Couplings and Resonances; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Saturn's Atmosphere; Saturn's Interior; Saturn's Magnetic Field; Saturn's Satellites.

Saturn's Satellites

Categories: Natural Planetary Satellites; Planets and Planetology; The Saturnian System

Saturn has a remarkably diverse set of satellites. They include gigantic Titan, which retains a thick atmosphere; Enceladus, possessing a vastly reworked surface that includes active geysers; Hyperion, a disk-shaped satellite whose rotation is erratic; Phoebe, moving in a retrograde orbit; and a coorbiting pair called Janus and Epimetheus, to name some of the most interesting of the sixty confirmed satellites.

OVERVIEW

Prior to the space age, Saturn was known as the beautiful ringed world of the solar system. Many of its numerous larger satellites were discovered prior to the time of interplanetary spacecraft, the most notable being Titan, Saturn's largest satellite and the only one known from telescopic observation to maintain a thick atmosphere. Prior to the Voyager flybys, planetary scientists expected all of the other Saturn satellites to be relatively uninteresting ice inactive worlds. Only Iapetus was a curiosity, since it displayed a very reflective side and an extremely dark side as well. The Voyager flyby results and the Cassini orbiter images and observations revealed Saturn's system to be a miniature solar system in its own right with a variety of extremely interesting and diverse satellites.

When Voyager 1 passed by Saturn's largest satellite Titan in November, 1980, scientists were somewhat disappointed with imagery transmitted back to Earth. Titan appeared as a uniform orange sphere whose outline was blurred by a dense cloud cover. Closer examination found a higher layer of ultraviolet haze. The southern hemisphere has a slightly darker cast than the northern hemisphere. A clear equatorial boundary was noted, and a darker polar ring is evident in some photographs from Voyager 2. Beneath those clouds, Titan proved more interesting. Voyager 1's close passage behind the disk of Titan allowed the use of its radio transmissions to probe the satellite's atmosphere. The pressure at ground level is 1.5 times that of Earth. If Titan's lower surface gravity is taken into account, the implication is that every square meter of Titan has ten times as much gas above its surface as Earth does.

Methane was spectroscopically detected from Earth, but the prime component of Titan's atmosphere proved to be nitrogen. It was suspected that as much as 10 percent of the atmosphere is argon, and methane makes up between 1 and 6 percent of the rest of the atmosphere, increasing in concentration near Titan's surface.

At higher altitudes, solar ultraviolet rays break methane down, and new molecules form as some hydrogen is lost. Spectroscopic observations show traces of hydrogen, ethane, propane, ethylene, diacetylene, hydrogen cyanide, carbon monoxide, and carbon dioxide. Together, these components form the petrochemical smog that so frustrated the Voyager imaging team.

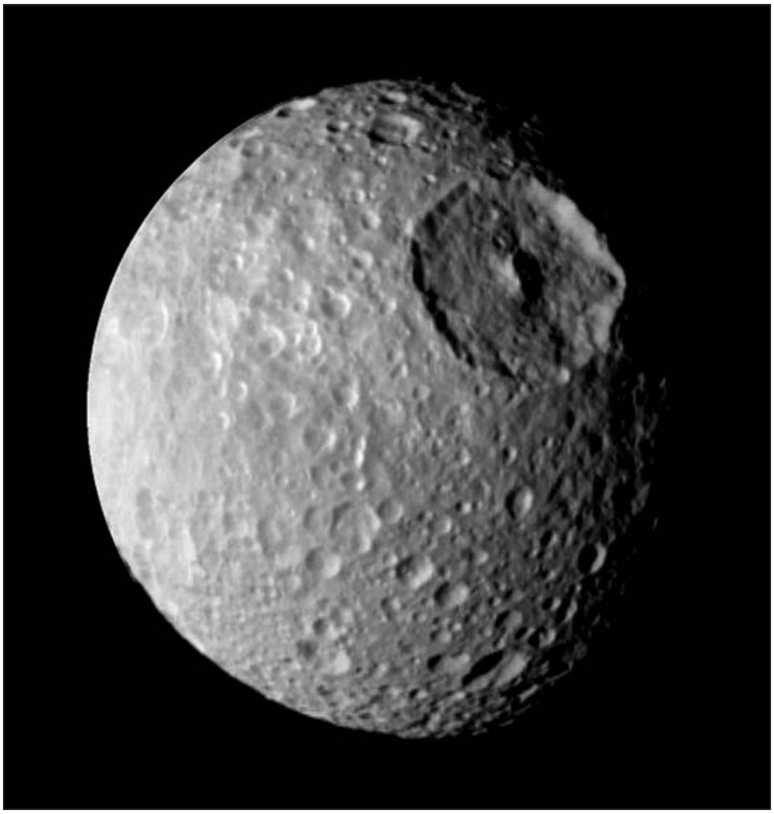
The upper optical haze layer lies about 280 kilometers above the surface. The main cloud deck is about 200 kilometers from the surface. Titan's solid surface is 400 kilometers smaller in diameter than previously thought, smaller than both Ganymede and Callisto in the Jupiter system. Why do these Jovian satellites not have atmospheres? Titan orbits at a greater distance from Saturn than either of these satellites do from Jupiter, so its tidal stress is less. Furthermore, Saturn is twice as far from the Sun as is Jupiter, so solar radiation intensity at Titan is four times weaker than in the Jovian system.

Beneath those tantalizing orange clouds, the surface temperature is only 94 kelvins; combined with the fact that Titan's atmospheric pressure is 1.5 bars, this temperature suggested the possibility of an ethane and/or methane sea on Titan's surface. If tidal stresses heat the interior enough, there may even be icy geysers. The possibility of life arising at such low temperatures appears unlikely, but certainly the carbon chemistry on Titan must be very interesting.

The European Space Agency (ESA) provided the Huygens probe, a combinations atmospheric entry probe and soft lander, for National Aeronautics and Space Administration's (NASA's)

Cassini program. The Orbiter carried Huygens from launch in October, 1997, to release on Christmas, 2004. For the next two weeks the probe flew independent from the Cassini orbiter, and then it entered Titan's atmosphere on January 14, 2005, and dropped down under a large parachute to a safe touchdown near Titan's Xanadu region. Some researchers expected Huygens to splash down in a cryogenic sea of liquid hydrocarbons. It became clear rather quickly that Huygens had "plopped" down in what some referred to as Titanian mud. Evidence of liquid action on the surface was found, but the original idea of liquid hydrocarbon seas were dashed. Analysis of data sent from the probe on the way down to impact revealed several layers in the atmosphere, most notably a thick haze between 18 and 20 kilometers above the surface. An Aerosol Collector and Pyrolyzer collected samples at different altitudes to determine the pressure of volatiles and organic materials. The probe was also outfitted with a gas chromatograph mass spectrometer to determine atmospheric composition. Titan's atmosphere proved to be hazier than expected, as dust particle concentration was greater than previously believed. Wind data suggested that Titan's atmosphere circulated gas from the south to north pole and back again in periodic fashion.

Winds would play a large role in planetary dynamics for this complex world. Indeed, two years after the probe's several hours of data were collected, Cassini scientists came to the conclusion that Titan's crust moves on a subsurface ocean with crust movements in part driven by wind actions. That movement was noted by comparing radar data from the Orbiter taken at different times during the mission in concert with available Huygens data. The proposed liquid subsurface layer would be located 50 to 100 kilometers beneath the crust and include liquid ammonia in a water ice slush. Floating on this layer, the crust was seen to move as much as 30 kilometers over the course of several years of Cassini observations.



Mimas, with its huge crater, Herschel. (NASA/JPL/Space Science Institute)

Further examples of Huygens and Cassini images eventually found the evidence proving the existence of ancient shorelines and the presence of liquid on the surface of Titan. Computer models of this dynamic world had to be greatly altered due to Huygens and Cassini data, and many new questions were raised to give Titan an even more mysterious nature. However, it still was believed to be a world rather similar in some ways to an early Earth, just a world frozen at an early point of physical and chemical evolution prior to the development of life. Down in the subsurface liquid layer higher temperatures could permit complex biochemistry, but there was no information produced by Cassini to investigate that supposition.

By measuring gravitational perturbations on the Voyager spacecraft as they flew through the Saturnian system, scientists at the Jet Propulsion Laboratory (JPL) could determine the masses and densities of Saturn's middle-sized satellites. Rhea's bulk density of 1.3 grams per cubic centimeter suggests that it contains more ice and fewer silicates than Titan. It is worth noting that the density of bodies depends not only on their composition but also on how tightly packed they are. The greater the mass, the higher the gravity, and thus the greater the density. Thus Titan, Ganymede, and Callisto, which all approximate the size of the planet Mercury, have very similar densities, about 1.9 grams per cubic centimeter, and the smaller, icy satellites are less dense, even though the composition may be quite similar to that of larger satellites.

Rhea and all the smaller Saturnian satellites lack atmospheres and show some signs of older, cratered surfaces. The trailing side of Rhea is covered with pale, wispy streaks, a type of feature it shares with Dione. These streaks may be evidence of venting of water vapor from these satellites' interiors, perhaps from tidally induced volcanism in the past. On Earth's moon, such activity was on the side facing Earth rather than on the trailing side. Perhaps these wisps were once found on Rhea's leading side but were eventually eroded, much as meteoric dust erases all but the youngest ray patterns around lunar craters.

During a Cassini flyby of Rhea in November, 2005, some surprising results were obtained. While the spacecraft's magnetometer did not pick up any interactions with Saturn's magnetosphere that would have indicated even a meager atmosphere about Rhea, there was evidence of a broad debris disk and one structured ring about this satellite. The debris disk extended several thousand kilometers out from Rhea, hence it was several Rhea radii in expanse. Computer simulations of the gravitational interactions of Saturn and Rhea indicated that this ring could exist for a considerable time. The source for the ring particles about the small satellite could have been a large impact event with Rhea.

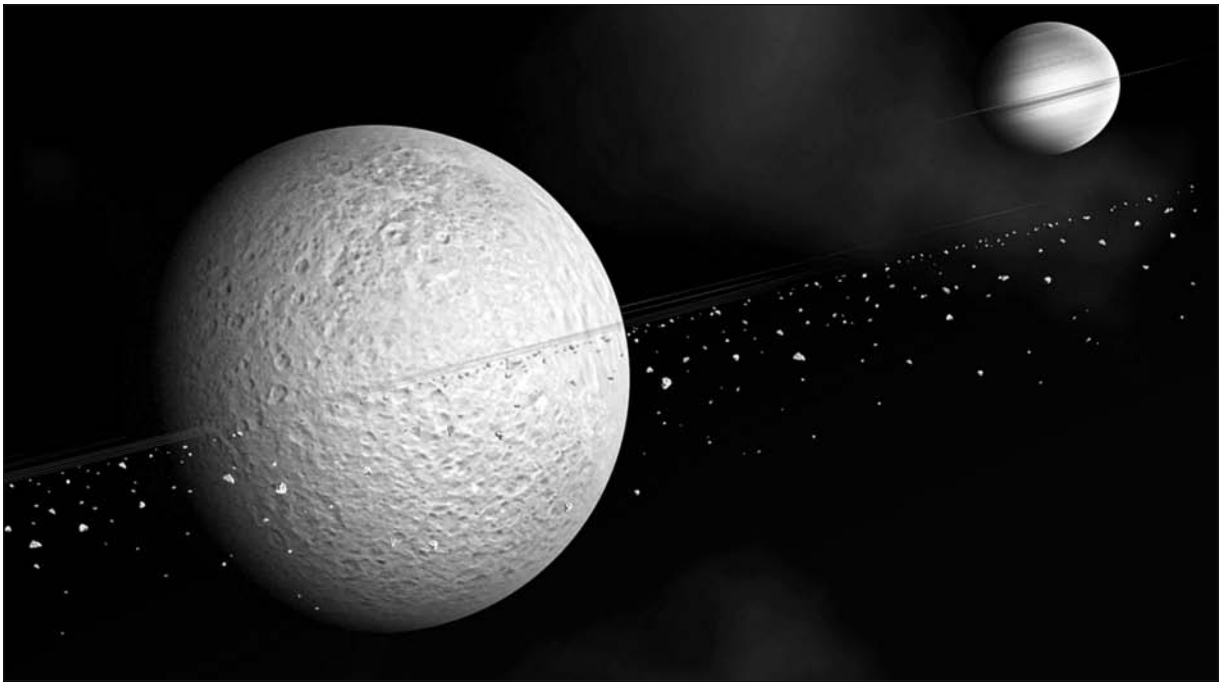
Slightly smaller than Rhea, Dione is bit more dense (1.4 grams per cubic centimeter). Its wispy patterns are more marked than those of Rhea, and it also has some long cracks and large areas of fairly fresh ice, without large craters in evidence. Ice may have flowed through cracks during the cooling phase of the satellite. Unlike most substances, water expands when it freezes. Therefore, satellites made mainly of ice, or differentiated with rocky cores and mantles composed largely of water, might show such expansion cracks. Such cracks are evident on Tethys, and were noted by Voyager 2 on the Uranian moons Ariel and Titania in January, 1986.

Tethys is similar in size to Dione, but it features one huge impact crater. The crater's floor is quite flat, suggesting internal flooding resulting from impact heating. Running from the crater three-quarters of the way around Tethys is a single gigantic valley system, Ithaca Chasma. Tethys's craters are lower in relief than lunar craters. The icy crust of these Saturnian satellites is more plastic than is lunar crust. Older terrain cratering appears to have been just as heavy as on Earth's moon, resulting craters appear less rugged than the lunar highlands.

The small satellite Mimas features a notable exception to the above rule. One huge crater, Herschel, is one-third as large as Mimas itself. This crater is very deep, about nine kilometers, with a central peak about six kilometers high. It constitutes one of the most striking geological features in the solar system. This crater on such a small roughly spherical body gives it the appearance of the Death Star station in the movie *Star Wars*. Many of the planetary scientists on the Voyager and Cassini teams, and astronomy professors world wide, fondly refer to Mimas as the "Death Star Moon."

It is likely that an impact of any greater force would have broken Mimas apart. Such huge impacts might also explain the very jumbled appearance of Uranus's Miranda, which was apparently broken into several large pieces, then haphazardly reassembled by gravity later.

One highly speculative hypothesis may account for such massive impacts and the intense cratering that is evident throughout the solar system. Perhaps a terrestrial planet in an un-



Saturn's second largest moon, Rhea, may have a debris ring of its own, as illustrated in this artist's rendition. (NASA/JPL/JHUAPL)

stable orbit beyond Mars and an outer icy Jovian satellite were totally fragmented in a high-energy head-on collision. The lighter, icy debris might account for some of the comets. Heavier chunks may have found a relatively stable orbit and formed the asteroid belt. Many chunks and particles, however, would have been scattered in all directions to impact other worlds. In some cases, the impact would have been forceful enough to send up debris which, in turn, would bombard neighboring worlds. This scenario might explain ices found on certain asteroids, the existence of captured satellites such as Phoebe and Phobos, and the fact that meteoritic material seems to have originated on a differentiated planet. It might also explain how fragments which scientists agree came from lunar basalts and from Mars are found on Earth.

While almost a twin of Mimas in size, and orbiting just beyond it in the satellite system, Enceladus is a very different world up close. Even the Pioneer 11 data indicated that it has an albedo near 100 percent. It seems to be made of fresher ice, reflecting far more sunlight than most Saturnian satellites. Had its material been older, dark meteoritic and cometary dust would have darkened it. Voyager's cameras revealed that one of its hemispheres is heavily cratered and fairly old, but the opposite side features smooth plains cut by grooved terrain, similar to Ganymede. This evidence of much rifting and recent internal activity appears on a satellite about one-tenth the size of Ganymede. A count of ring particles in Saturn's extended E ring also found that they peaked near Enceladus. Just as the dust ring of Jupiter is supplied by Io's volcanoes, icy geysers on Enceladus periodically shoot debris above this active world. Why is this small world active at all? Mimas lies closer to Saturn and thus is more tidally stressed, yet it shows no such activity. Nor does the proximity of any other large satellite seem to account for the heating required to generate such activity. Such activity on smaller satellites is not unique; Uranus's moon, Ariel, which is similar in size to Mimas, shows obvious broad rift valleys. The source of heating for this extensive and possibly continuing crustal activity is a mystery.

Cassini found geysers on Enceladus near its south pole along long cracks which essentially act like vents. Fresh crystalline ice forms at the site of these cracks and colors the features distinctively. Cassini scientists dubbed these nearly parallel cracks found at the south polar region "Tiger Stripes." The Tiger Stripes were found to be 124 kilometers long and 40 kilometers apart. This activity was not new, so why did the Voyagers fail to see these geysers? Voyager 2 flew over Enceladus's north pole and missed them. Cassini's near-infrared mapping spectrometer and solid state imager both examined the ice around the Tiger Stripes. Freshly formed ice was crystalline. As time progresses that pristine ice becomes radiation-damaged amorphous ice.

Data and the geyser actions strongly suggested the presence of a subsurface ocean on Enceladus. However, calculations about heat transport inside Enceladus led researchers to believe that the satellite's subsurface ocean would not be able to exist for more than 30 million years if it were warmed only by heat escaping from the core toward the crust. Since that ocean most likely has been in existence for more than 30 million years, the heating mechanism for both the subsurface and the cryovolcanic activity at the satellite's south pole must be from tidal flexing. Only that could provide the 5.8 gigawatts of heat Cassini saw emerging from the Tiger Stripes over which it flew on a close fly. Since internal heat sources apart from tidal flexing produce only 0.32 gigawatts, without tidal heating Enceladus's subsurface ocean would have frozen. But the story here is more complex than that. Without the subsurface ocean, tidal flexing of the magnitude necessary to produce the observed heat would not be possible, and without the heat for the tidal flexing the ocean would freeze.

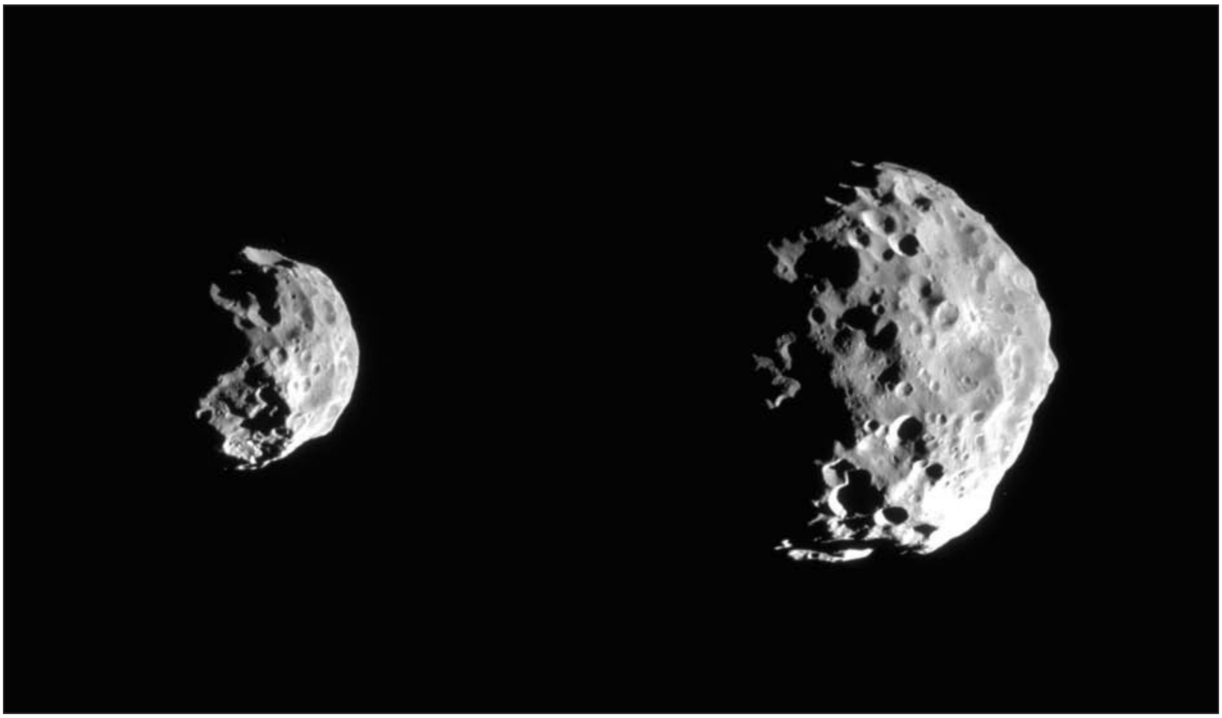
Iapetus confronts scientists with another mystery. A portion of this satellite is extremely dark, whereas the rest of Iapetus has an albedo typical of an icy surface. Iapetus's dark side is six times lower in albedo than that icy portion. Within the darker portion is an irregular dark spot. This pattern of darker leading hemispheres is also seen on Rhea and Dione, but to a far lesser degree. The brighter, icy side does have an albedo of 50 percent. This is typical of older water-ice crusts, and it shows heavy cratering typical of other similar satellites. Some larger craters near the boundary between the hemispheres have light-colored walls, with darker flat floors, like some larger craters on Earth's moon. What is the reddish-black material that gives the darker side an albedo of only 5 percent? It is probably an organic tar, and it appears to be a good match with carbonaceous chondrite meteorites, the dark rings of Uranus, and the black crust of Halley's comet. Carbon almost always appears in an oxidized form (carbon dioxide, carbonic acid, carbonate rocks, carbohydrates) in the inner solar system. Dark neutral carbon is a major solid material found in the outer solar system. It was impossible to judge the age of Iapetus' dark spot, as Voyager's cameras could not pick up any details on the dark side. In fact, some photographs even make the dark side disappear into the blackness of space. The concentration of the dark material on the leading side suggested an external source for this coating. It had been hypothesized that dark Phoebe was responsible. Phoebe's color, however, does not match the black side of Iapetus. The dark floor of some of Iapetus's craters constitutes evidence for an internal origin. New data and insights would have to wait for Cassini to pass by Iapetus at a much closer distance than had the Voyagers.

Cassini flew within 1640 kilometers of Iapetus on September 10, 2007. Unfortunately during the encounter, Cassini entered a safe mode after on-board delicate solid state electronics suffered a cosmic ray hit. Fortunately, most of the science harvest was recovered after a short delay in playback. Among the findings was a raised area around the satellite's midsection that gave Iapetus the appearance of a walnut. Why the equatorial bulge on this unusual satellite? Julie Castillo of the Jet Propulsion Laboratory advanced an innovative explanation for the unique feature. Castillo invoked a high rotational speed early in the satellite's history coupled with heat from internal radioactivity, perhaps from aluminum 26 and iron 60 isotopes, that softened the satellite to form the equatorial bulge. Consideration of the time frame in which tidal forces forced the spin rate to diminish led to the conclusion that this particular pair of isotopes would be required, since they would be abundant and would generate heat quickly due to rapid radioactive decay. Then, from a softened and malleable state, the satellite's bulge was frozen in place before Iapetus's spin rate slowed down. More investigation would be needed to confirm or refute this theory. Unless extended mission priorities are changed and trajectories reevaluated, this could be the closest Cassini would ever get to this highly unusual satellite.

Saturn's outermost known satellite Phoebe may be a captured asteroid from the far reaches of the main belt, and therefore similar to Chiron. Phoebe's orbit is retrograde, like those of Jupiter's four outermost moonlets. All these small worlds are quite distant from the gas giants whose gravity trapped them. Cassini encountered Phoebe on the way into Saturn orbit insertion. Photographs revealed Phoebe's surface to look almost spongelike, not at all like the other Saturnian satellites.

Like the dark side of Iapetus, Phoebe has an albedo of about 5 percent, which is similar to those of two other captured asteroids, Deimos and Phobos, which orbit Mars. At 200 kilometers in diameter, Phoebe is much rounder than the odd "hamburger moon," Hyperion, which orbits between Tethys and Iapetus. Puck, a satellite of Uranus, is similarly round and dark, and about the same size as Phoebe. This fact suggests that a round shape is the norm for dark, primitive bodies such as these, and that something unusual happened with Hyperion.

Hyperion's shape is quite striking. It is a huge disk, about 250 kilometers across but only 150 kilometers thick. Like Phoebe, its surface is dark, old, and heavily cratered. Stranger still is Hyperion's rotation period. It has not yet been well defined. Like Earth's moon, Saturn's other satellites are tidally locked, with one side permanently facing the planet. As the Voyager 2 team tried to orient photographs of Hyperion to map it, however, they found that it was rotating chaotically. It appears to have no regular rotation period; it tumbles irregularly. Close coupling between Hyperion's eccentric orbit and that of Titan may cause this unique effect.



Phoebe, outermost known satellite, with its heavily cratered surface, may be a captured asteroid. Two images of the satellite shown here were imaged from the Sun-Phoebe spacecraft. (NASA/JPL/Space Science Institute)

The nine satellites discussed to this point were known well before the Voyager missions, but Voyager photographs found or confirmed eight more satellites, making Saturn's the most numerous satellite system. Several photographs suggested the existence of even more Saturnian satellites, but their periods of revolution and orbits had to be determined before they were formally recognized. All these new satellites are much closer to Saturn than are Phoebe and Iapetus, and they show a much more reflective, icy surface like Enceladus. None of these satellites is large enough to be nearly spherical or to have become differentiated. All of them have quite interesting orbits.

Just as the Trojan asteroids share Jupiter's orbit, so two of Saturn's middle-sized satellites have smaller companions in their orbits. Dione has two; Helene, the leading one, appears quite elongated, while the following one appears rounder. Lagrangian, Tethys's companion, is a smaller version of Mimas, with a huge crater from an impact that almost destroyed it.

Saturn's coorbital satellites were first spotted in 1966, but for more than a decade thereafter they were mistaken for a single satellite with an orbit under that of Mimas. Even prior to the Voyager flights, however, observers repeatedly noticed inconsistencies that led some to argue that there must be two satellites sharing the same orbit. Janus, the larger, is about 200 kilometers across, and Epimetheus is about 150 kilometers across. Actually, their orbits are not quite identical. The inner satellite has a period of 16.664 hours; the outer one has a period of 16.672 hours, or a difference of 29 seconds per orbit. Every four years, the inner satellite overtakes the outer at the speed of nine meters per second, and they exchange orbits. This close relationship and the irregular, elongated appearances of these satellites suggest they were once part of a single larger one split apart by a collision into the two pieces now sharing the same orbit.

The inner three satellites discovered by Voyagers 1 and 2 are all closely associated with Saturn's rings. Atlas, a tiny, football-shaped body, orbits just outside the bright A ring of Saturn. Prometheus is a shepherding moon, keeping the particles in Saturn's F ring in place from the inside of that ring. Pandora plays a similar role on the outside of the F ring. Their close relationship to this thin set of ringlets may explain why the F ring sometimes appears braided. Additional satellites were identified in subsequent reviews of Voyager and other available data. Then, with the arrival of the Cassini probe in the Saturn system, the number of recognized satellites again increased significantly, reaching 60 by 2008.

KNOWLEDGE GAINED

While Jupiter possesses four satellites comparable in size to Earth's moon or even to Mercury, Saturn has only one, Titan. Like Jupiter's Ganymede and Callisto, Titan is comparable to Mercury in size but only about one-third as massive and dense. Its exact dimensions were still in debate prior to the Voyager missions. Its visible orange disk made it appear to be the largest known natural satellite; out-of-date astronomy textbooks will list Titan as the largest satellite in the solar system. Spectroscopic observations plainly revealed an atmosphere with gaseous methane and other hydrocarbons. Just how deep was the atmosphere, and what was it made of? These questions led the Voyager 1 team to target Titan as a main mission objective and to guide one probe closer to this satellite than to any other body on its mission.

The chief discoveries that resulted concerned Titan's atmosphere. It is thick, twice as dense as Earth's, but like Earth's, Titan's atmosphere is made primarily of nitrogen. Orange clouds appear to be a hydrocarbon smog, with complex organic chemistry taking place there. Surface temperatures and pressures lie close to the triple point of methane, so the surface might experience methane rains that would build up into lakes of liquid methane and freeze into methane ice at the poles. Confirmation of that would have to wait for a probe outfitted with imaging radar and/or a lander. Thus the origin of the Cassini mission. Combination of radar images taken from orbit with data from the Huygens lander eventually confirmed the presence of cryogenic lakes and found ancient shorelines of lakes no longer existent. Huygens appeared to have landed in a wet slushlike material at cryogenic temperatures rather than floating on a lake or sea or even having hit a hard icy surface.

By the time its primary mission was completed, Cassini had flown past Titan several dozen times at varying distances. Perhaps one of Cassini's biggest surprises was the detection that the surface of Titan moved as much as 30 kilometers between the earliest flyby of the Cassini primary mission (2004) and some near the time that the extended mission was approved (2008). This suggested that the crust floated on a layer of fluid, meaning the large satellite likely has an underground ocean, presumably a mixture of water and ammonia.

All of Saturn's remaining satellites are smaller than Earth's moon. Rhea is next in size, about half as large as the Moon at 1,500 kilometers in diameter; Voyager 1 showed its icy surface to be cratered, but with fresher ice creating wispy terrain. Dione is next in size, at 1,100 kilometers in diameter, and has even more wispy terrain than Rhea. Tethys featured a huge, flattened crater on one side, with a great crack or rift running to the other side.

The innermost of the satellites well-known prior to the Voyager flybys are Mimas and Enceladus, both about 500 kilometers in diameter. Mimas was found by Voyager 1 to have a dramatic impact crater one-third as large as the satellite. Enceladus is one of the most puzzling satellites, with tidal stresses producing plate activity, according to Voyager 2 data. These satellites are, in order from Titan inward: Dione, Tethys, Rhea, Enceladus, and Mimas. Prior to the Voyagers, only their orbital periods and approximate diameters were known, based on their brightness. No one had actually seen their disks. Cassini provided data that made Enceladus much more interesting to planetary scientists.

The brightness of Iapetus presented a major problem. When Gian Domenico Cassini found it in 1671, he realized that this odd satellite must be far brighter on one side (the leading hemisphere as it orbits Saturn) than on the other. Diameter measurements were impossible until Voyager photographed the disk. It proved to be about half as big as the Moon, with one side bright and icy. The other side is mostly covered with a layer of tarlike black material that hid any surface features from the cameras on Voyager 2. Similarly, little was known about Hyperion, another dark satellite orbiting between Titan and Iapetus; it was found by Voyager 2 to be irregularly shaped and tumbling without any rotational period.

Phoebe, the outermost known satellite, is distinguished by its retrograde orbit, like four of the outermost satellites of Jupiter. Like the dark side of Iapetus, Phoebe may be covered with carbon-rich material. More puzzling, the existence of Janus, a tenth moon, had been suspected, but before Voyager, photographs showed it in the wrong place. Voyager 1 detected two satellites sharing the same orbit. The other seven of Saturn's major satellites were not known prior to the Voyager missions. Cassini added considerably to the total list of Saturn's family of satellites.

CONTEXT

Practically nothing was known about Saturn's satellites prior to the Voyager flybys. Titan, Dione, Mimas, and Rhea were examined most fully by Voyager 1, in November, 1980. Until the arrival of Cassini in Saturn orbit, most information about Enceladus, Iapetus, Hyperion, and Tethys had come from Voyager 2 in August, 1981. Much about these satellites were discovered or confirmed by Voyager 1, but thanks to improved orbital data, they were best photographed by Voyager 2. Clearly, a strong argument can be made for using two spacecraft in flyby missions.

In brief, the Voyager missions found Saturn's satellite family to be a very diverse lot. Even satellites similar in size and mass, such as Mimas and Enceladus, appeared very different up close, and obviously were shaped by different processes. Each satellite has its own history of impacts. Tidal stress has played an important role in the evolution of many of these bodies, as it has in the Jovian satellite system. Each satellite has its own fascinating evolutionary story to be interpreted by geologists.

With Cassini repeatedly orbiting Saturn and conducting numerous flybys of many of the satellites, planetary scientists were able to make comparisons over time. Just as the Voyagers had piqued interest in satellites that had once been thought to be merely crater-pocked ice balls, Cassini images revealed many of the satellites not well studied by the Voyagers to also be rather intriguing in totally unexpected ways. Interest in Endeladus, for example, increased greatly due to Cassini observations.

J. Wayne Wooten and David G. Fisher

FURTHER READING

Consolmagno, Guy. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994. A text accessible to college-level science and nonscience readers alike. Presents subjects at low-level mathematics and also involves integral calculus where required. Demonstrates how the area of planetary science progresses by questioning previous understanding in light of new observations.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that have investigated all planets with the exception of Pluto by the publication date. Takes an astrophysical approach to give our solar system a wider context as just one member of similar systems throughout the universe.

Harland, David M. *Cassini at Saturn: Huygens Results*. New York: Springer, 2007. This text provides a thorough explanation of the entire Cassini program, including the Huygens landing on Saturn's largest satellite. Essentially a complete collection of NASA releases from the start of Cassini flight operations through the majority of Cassini's seventy orbits of its primary mission. Cassini's primary mission concluded a year after this book entered print. Technical writing style but accessible to a wide audience.

_______. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002. Another book in Springer's Space Exploration Series, this is a technical description of the Cassini program, its science goals and the instruments used to accomplishment those goals. Written before Cassini arrived at Saturn. Provides a historical review of pre-Cassini knowledge of the Saturn system.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text on planetary science. The chapter on Saturn covers all aspects of ground-based and spacecraft observations of Saturn.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Available to the serious general audience.

Leverington, David. *Babylon to Voyager and Beyond: A History of Planetary Astronomy*. New York: Cambridge University Press, 2003. An historical approach to planetary science. Heavily illustrated, concludes with a summary of spacecraft discoveries. Suitable for general readers and the astronomy community.

Lorenz, Ralph, and Jacqueline Mitton. *Lifting Titan's Veil: Exploring the Giant Moon of Saturn*. Cambridge, England: Cambridge University Press, 2002. An in-depth examination of all that was known about Titan prior to the Cassini-Huygens mission written by an engineer who worked for the European Space Agency and an astrophysicist who was Press Officer for the Royal Astronomical Society. Describes the mission of Cassini-Huygens, but the book was published before the spacecraft arrived at Saturn.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/ Addison-Wesley, 2003. A fine survey of the solar system, and very current for the editing date. Intended for use as a introductory text, it is very well organized. Highly recommended for the general reader.

Van Pelt, Michel. *Space Invaders: How Robotic Spacecraft Explore the Solar System*. New York: Springer, 2006. An historical account of robotic planetary science missions attempted by all spacefaring nations written by an European Space Agency cost and systems engineer. As such the narrative not only explains the science but also provides a behind-

the-scenes description of the development of a space exploration mission from concept proposal to flight operation.

See also: Enceladus; Iapetus; Io; Jovian Planets; Jupiter's Ring System; Jupiter's Satellites; Neptune's Ring System; Neptune's Satellites; Planetary Ring Systems; Planetary Satellites; Saturn's Ring System; Titan; Uranus's Rings; Uranus's Satellites.

Search for Extraterrestrial Intelligence

Category: Life in the Solar System

The search for intelligent life in the universe is perhaps the most profound of human endeavors. If extraterrestrial intelligence were discovered and communication established, it would irrevocably alter the conception of humanity's place in the universe. Contemporary science suggests that given a suitable planetary environment and sufficient time, life will evolve, and since intelligence and technology have high survival value, they will inevitably follow. Given that there are a myriad of suitable planets in the galaxy, intelligent life-forms, willing and able to communicate, should be abundant.

OVERVIEW

The search for extraterrestrial intelligence is a relatively recent exploratory science which assumes that if intelligent life, more technologically advanced than us, exists elsewhere in the universe, they will attempt to communicate by broadcasting messages using radio waves. Detecting and decoding such messages, however, requires considerable effort. If there is only a vanishingly small probability that an alien civilization has evolved, it would be a futile waste of time and money to search for a message. In order to arrive at an estimate of the number of technological civilizations with the ability and desire to communicate, Dr. Frank Drake convened a small group of scientific experts from various fields at the National Radio Astronomy Observatory (Green Bank, West Virginia) in 1961. This group formulated a simple equation to estimate the possible number of communicative civilizations. The equation is:

$$N = R^* \, f_p \, n_e \, f_l \, f_i \, f_c \, L$$

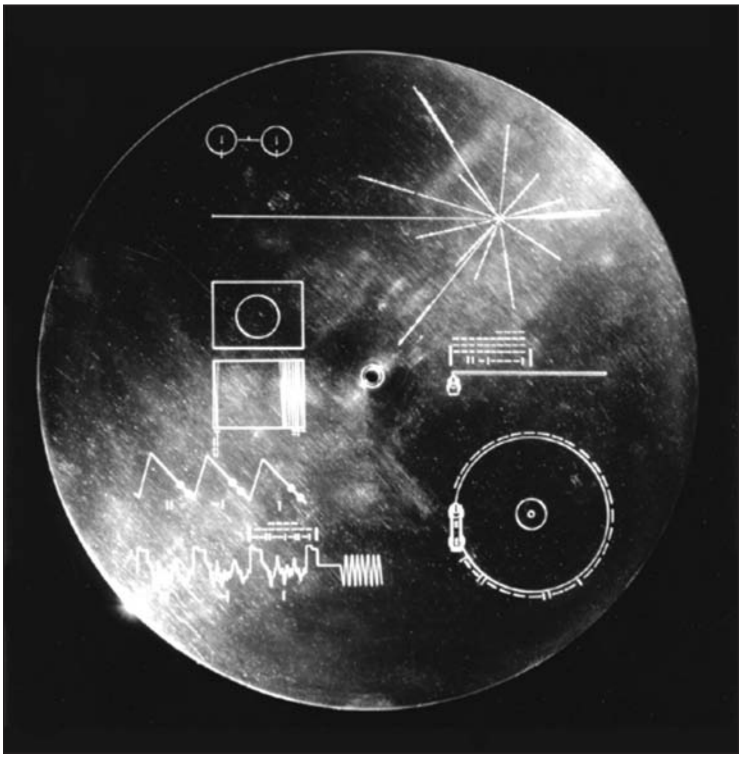
where N is the number of advanced technological civilizations, R* represents the rate of formation of suitable stars, f_p stands for the fraction of stars with planetary systems, n_e symbolizes the number of planets in a solar system which are suitable for life, f_l corresponds to the fraction of suitable planets on which life evolves, f_i signifies the fraction of inhabited planets on which intelligent life evolves, f_c characterizes the fraction of planets with intelligence on which a technological civilization can emerge, and L embodies the length of time the technological civilization would broadcast signals. When these factors are multiplied, the equation will yield an estimate of N.

After considerable discussion, the group narrowed the possible range of values for each factor to the following values. The average rate of star formation is the 100 billion stars in our galaxy divided by its 10 billion year age, yielding about 10 stars per year. The fraction of stars with planetary systems was estimated to be about one, but the number of habitable planets was thought to be only about one in ten (0.1). Since it is believed that life evolves wherever the conditions are conducive, the fraction of habitable planets on which life evolves was taken to be one. Because intelligence has high survival value, it was assumed that the fraction of life-bearing planets on which intelligence evolves is also one. The fraction of planets on which intelligent life develops technology is taken to be one because of technology's high survival value, and from the fact that tool-using cultures arose independently at multiple locations on Earth.

The final factor, the average lifetime of a technological civilization, is the most difficult to estimate. The Earth has had the ability to communicate by radio waves for less than a hundred years, and this state may not last another hundred years if the environmental and social problems which plague Earth are not addressed and solved. If an advanced civilization can continue in its communicative state for at least a thousand years, it could probably last for a million years. When the foregoing factors, momentarily excluding L, are multiplied, the result is a value of about one. Thus, if a technological civilization exists for only 200 years, the number of communicators in our galaxy will be only about 200 out of the 100 billion stars present; they would be few and far between. If, on the other hand, advanced civilizations can exist for at least a million years, there will be enough potential communicators to make the search worthwhile. Scientists working on the Search for Extra-terrestrial Intelligence (SETI) project have chosen the optimistic view.

The SETI Institute (Mountain View, California), was founded in 1984 as a private, nonprofit organization dedicated to searching the skies for technological indicators of extraterrestrial intelligence. Because of the vast distance between stars, the fact that no material object can travel faster than the speed of light, and the prohibitive cost of interstellar space travel, it is assumed that an alien civilization would attempt to locate other intelligent species by using radio waves to carry nonrandom messages. Two problems then present themselves. Which stars are likely candidates for intelligent life to evolve, and what radio frequencies should be scanned for possible messages? The stars most likely to have habitable zones where planets would be neither too hot nor too cold for life are main sequence G- and K-type stars. G-type stars, such as our Sun, have a surface temperature of about 6,400 kelvins and a habitable zone that extends from Venus through Mars. K-type stars have a surface temperature of approximately 4,250 kelvins and a somewhat narrower habitable zone. Stars hotter than G probably evolve too quickly for life to develop intelligence, while stars cooler than K would have such a narrow habitable zone that a planet is unlikely to occupy it.

Only certain radio frequencies are useful for communication. From Earth's surface, the available window of frequencies ranges from 1 gigahertz (GHz) to about 30 GHz. Frequencies lower than 1 GHz would be lost in the background radio noise of the Galaxy; frequencies higher than 30 GHz are absorbed by Earth's atmosphere. There are still far too many possible frequencies available in this window to make a search practical. There are, however, two important frequencies within this window that would be likely choices for an advanced civilization, wishing to establish contact with an alien civilization, to broadcast a message. These are the 14.3 GHz frequency, radiated by neutral hydrogen (H), and the 16.7 GHz frequency, emitted by the



The two Voyager spacecraft carried identical copies of this gold-plated copper record, which was designed to tell the story of Earth to other intelligent civilizations that might find the spacecraft. Data included music, greetings in sixty languages, sounds from Earth, diagrams, and photographs. (NASA) hydroxyl ion (OH). Because hydrogen is the most abundant element in the universe and because H and OH combine to form water (HOH, or H_2O), any other intelligent species based on water would be reasonably likely to choose one of these frequencies on which to broadcast. Although it is possible that an alien life-form could be formed from other molecules, there are valid chemical reasons to assume that this is unlikely.

KNOWLEDGE GAINED

No direct evidence of extraterrestrial life has ever been observed, but there exists some significant data related to the origin of life. Complex carbon molecules, the precursors of life, have been found in interstellar clouds and in certain meteorites from our solar system. Based on the not unreasonable assumption that intelligent life is abundant in our galaxy, the National Aeronautics and Space Administration (NASA) funded a search for alien radio messages during the late 1980's. This project was canceled in the early 1990's because politicians feared public ridicule for funding a search for "little green men." Although the annual cost was small (about the same as one Air Force attack helicopter) it was felt that the money would be better spent elsewhere. The SETI Institute continued the search using the radio telescopes at major radio observatories. Stars deemed likely to have planets on which intelligent life could evolve are monitored for nonrandom radio emissions; to date none has been identified.

A joint effort between the SETI Institute and the Radio Astronomy Laboratory at the University of California, Berkeley, is the Allen Telescope Array (ATA), a composite of 350 separate radio dish antennas at one location in northeast California. Designed for innovative astronomical research, the array is also being used to search for nonrandom extraterrestrial signals. Because the ATA is equivalent to one huge radio telescope, it can collect enormous amounts of data every hour of the day, every day of the year. It can also scan a wider portion of the sky and do so more quickly than existing telescopes. The ATA is programmed to survey one million stars within 1,000 light-years for extraterrestrial signals in the frequency range of 1 to 10 GHz. It will also survey the billions of stars of the inner galactic plane in the frequency range from 14.3 GHz through 16.7 GHz.

Another innovative experiment searches the sky for intelligent signals in the form of powerful pulsating flashes of laser light beamed from solar systems many light-years distant. The new pulse-detection system, coupled to the Lick Observatory's 40-inch telescope, is able to detect beacons pulsating at less than a billionth of a second with an error rate of only about one per year. This system can be automated and promises results less ambiguous than previous optical systems in which false alarms occurred daily.

CONTEXT

From the ancient Greek concept that Earth was the center of the universe, humans have always assumed their place in the cosmos to be pivotal and unique. As science advanced in the sixteenth and seventeenth centuries, Nicolaus Copernicus and Johannes Kepler showed that the Sun was the center of the known universe and Earth was merely another planet. In the early twentieth century, Edwin Hubble discovered that the Sun is a very common G-type star in a galaxy of a hundred billion stars and that the Milky Way galaxy is only one among billions in the universe.

There is nothing particularly unique about the Sun. Many humans, however, clung to the notion that few other stars were likely to have Earth-like planets. By the mid-twentieth century, however, astronomers had concluded that such planetary systems would be quite common. In order to preserve the illusion of uniqueness, it was then assumed that life on Earth was unique. Discoveries have shown, however, that the molecular precursors of life, such as amino acids, have formed elsewhere in the solar system, leading biologists to conclude that life is probably abundant in the universe. Since intelligence has a high survival value, it follows that intelligent life would not be uncommon either. Because a civilization would necessarily be more advanced to send an intense signal through interstellar space, humans would also be denied their unique rank as the supreme intelligence. Confirmation of a superior extraterrestrial intelligence would be the kiss of death to the lingering remnants of the unjustifiably self-centered sense of our uniqueness and importance in the universe.

George R. Plitnik

FURTHER READING

Cameron, A. G. W., ed. *Interstellar Communication*. New York: W. A. Benjamin, 1963. One of the earliest books to present a serious consideration of the possibilities and ramifications of communicating with extraterrestrials, this collection of reprints and original articles covers all the extant knowledge available at the time of publication.

Chaisson, E., and S. McMillan. *Astronomy: A Beginner's Guide to the Universe*. 5th ed. Upper Saddle River, N.J.: Pearson/Prentice Hall, 2007. This accessible work, written for laypersons with inquisitive minds, has an entire chapter devoted to life in the universe. Commencing with the nature and origin of life on Earth, it first considers the possibility of life elsewhere in our solar system and then details the best estimates of how many communicative civilizations may exist in our galaxy. Finally, the various searches being conducted for extraterrestrial intelligence are discussed.

Christian, James L. *Extraterrestrial Intelligence: The First Encounter*. Buffalo, N.Y.: Prometheus, 1976. A fascinating speculative account based on the not unreasonable assumption that extraterrestrial life is abundant in our galaxy and able and willing to communicate. Fourteen authors consider what might be humanity's response, in terms of our view of our place in the universe and how our theologies may respond, when contact is finally made.

Ponnamperuma, C., and A. G. W. Cameron. *Interstellar Communication: Scientific Perspectives*. Boston: Houghton Mifflin, 1974. Covers every aspect of interstellar communication from a scientific perspective. Various experts discuss the prevalence of planetary systems in our galaxy, the likelihood of communicative intelligences evolving on other planets, methods of communicating, and interstellar probes.

Seeds, Michael A. *Foundations of Astronomy*. 9th ed. Belmont, Calif.: Thomson Brooks-Cole, 2007. This lavishly illustrated text commingles experimental evidence and theory to provide deep but well-explained elucidations of many fascinating facets of the universe. The final chapter, "Life on Other Worlds," is a comprehensive survey of the nature of life, its origin on Earth, life in our solar system, life on other planetary systems, communicating with distant civilizations, and estimating the number of technological civilizations in our galaxy.

Shklovskii, I. S., and Carl Sagan. *Intelligent Life in the Universe*. San Francisco: Holden-Day, 1966. The definitive tome covering the entire panorama of the natural evolution of the universe, including the development of intelligence and technical civilizations in our galaxy. Of the thirty-five chapters, eleven deal with life in the universe; the final chapter discusses intelligent life in the universe in great detail.

See also: Extraterrestrial Life in the Solar System; Habitable Zones; Life's Origins; Main Sequence Stars; Mars: Possible Life; Radio Astronomy.

Solar Chromosphere

Category: The Sun

The solar chromosphere is the layer of the Sun's atmosphere, a few thousand kilometers thick, immediately above the photosphere. The gas is warmer, thinner, and more transparent than in the photosphere. The chromosphere is most directly visible from Earth as a layer of color during a total solar eclipse. Its spectrum shows bright emission lines, unlike the dark absorption lines of the photosphere. The element helium was first discovered in the Sun's chromosphere during a total solar eclipse.

OVERVIEW

The Sun is a ball of gas without a solid surface. The Sun's photosphere, however, is usually considered to mark the surface of the Sun. Relatively opaque, the photosphere blocks our view of the solar interior. It is also the coolest layer of the Sun. When astronomers observe the solar spectrum, they see dark absorption lines from the photosphere because the photosphere is cooler than the hot solar interior.

The chromosphere is the layer of the Sun's atmosphere directly above the photosphere. The chromosphere is only a few thousand kilometers thick. The temperature of the gas in the chromosphere is slightly higher than that in the photosphere. The bottom layer of the photosphere is at 6,600 kelvins, but the coolest level at the top of the photosphere is at 4,400 kelvins. For reasons that are not completely understood, the temperature in the chromosphere climbs from this low to about 6,000 kelvins as the distance from the photosphere increases. The chromosphere has a lower density than the photosphere; its density decreases with height above the photosphere from nearly 10^{-5} kilogram/meter^3 at the photosphere-chromosphere boundary to about 10^{-10} kilograms/meter^3 in the transition region between the chromosphere and the corona.

The layer of the Sun's atmosphere above the chromosphere is the corona. In the approximately 100-kilometer-thick transition region between the chromosphere and corona, the temperature rapidly increases with height above the photosphere, from about 6,000 kelvins to a few hundred thousand kelvins. The temperature in the corona can be millions of kelvins, but the gas is very thin. In the transition region, the density drops rapidly to about 10^{-12} kilograms/meter^3.

Because the chromosphere is so thin and transparent compared to the photosphere, it cannot be seen when the photosphere is visible. The much brighter photosphere overwhelms the fainter chromosphere. The chromosphere is briefly visible during total solar eclipses. When the Moon blocks light from the photosphere, the chromosphere shines briefly until the Moon also blocks the chromosphere. The chromosphere is visible only briefly, just before and just after totality. It is also possible to

Facts About the Sun

	Sun	*Earth*
Mass (10^{24} kg)	1,989,100	5.9742
Volume (10^{12} km^3)	1,412,000	1.083
Volumetric mean radius (km)	696,000	6,371
Mean density (kg/m^3)	1,408	5,515
Surface gravity at equator (m/s^2)	274.0	9.80
Escape velocity (km/s)	617.7	11.2
Ellipticity (oblateness)	0.00005	0.00335
Absolute magnitude	+4.83	—
Luminosity (10^{24} J/s)	384.6	—
Mean distance from Earth (10^6 km)	149.6	—
The Sun's Atmosphere		
Surface gas pressure (top of photosphere)	0.868 millibar	
Effective temperature	5,778 kelvins	
Temperature at bottom of photosphere	6,600 kelvins	
Temperature at top of photosphere	4,400 kelvins	
Temperature at top of chromosphere	~30,000 kelvins	
Photosphere thickness	~400 km	
Chromosphere thickness	~2,500 km	
Sunspot cycle	11.4 yrs.	
Photosphere composition:		
Hydrogen	90.965%	
Helium	8.889%	
Oxygen	774 ppm	
Carbon	330 ppm	
Neon	112 ppm	
Nitrogen	102 ppm	
Iron	43 ppm	
Magnesium	35 ppm	
Silicon	32 ppm	
Sulfur	15 ppm	

Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.

see the chromosphere using a coronograph, which is a disk in the focal plane of the telescope that blocks the photosphere to reveal the chromosphere and corona.

A dark, or absorption line, spectrum is produced when light from a hot, compressed gas passes through a cooler, thin gas. Therefore, the Sun's photosphere produces an absorption line spectrum. A bright, or emission line, spectrum results from a hot, thin gas. When astronomers observe the Sun's chromospheric spectrum, they are observing the chromosphere at the edge of the Sun, so the hotter interior is not directly behind the chromosphere in their line of sight. Hence, they observe emission lines. These emission lines flash into view for a short time just before and after the total portion of the eclipse. The chromospheric spectrum thus revealed is therefore called the flash spectrum.

The Sun is mostly hydrogen, and the brightest emission line in the chromospheric spectrum is the red hydrogen H-alpha line. This red emission line gives the chromosphere its reddish-pinkish color. The color also gives the chromosphere its name; *chromo* comes from the Greek word for color. Two spectral lines from the element calcium, the H and K lines, are also prominent in the chromospheric spectrum.

Astronomers use the H-alpha line and the H and K lines of calcium to observe the surface of the chromosphere. The chromosphere is relatively transparent at most wavelengths, but it is fairly opaque at the wavelengths of these spectral lines. Hence, observing the Sun at these wavelengths allows astronomers to image the chromosphere rather than the photosphere. The light at these wavelengths originates from different depths in the chromosphere, so comparing images made at the different wavelengths gives astronomers a three-dimensional image of the Sun's surface. Lyot filters made of calcite crystals are best for this purpose, but they are very expensive. Interference filters are less expensive, and H-alpha interference filters are available at a reasonable cost to allow amateur astronomers to observe the Sun safely with small telescopes.

Just below the photosphere, the Sun transfers energy from the interior via convection currents. These convection currents produce structures on the surface of the photosphere, called granules, that are the tops of the convection current cells. Faculae are bright areas on the solar photosphere that are often found in the lower areas marking the boundaries between the photospheric granules. When faculae extend upward from the photosphere into the chromosphere, they are called plages. Plages are therefore brighter than normal regions of the Sun's chromosphere. Plage regions can have opposite magnetic polarities, so magnetic field lines flow outward from one plage and connect to another plage where they flow back into the Sun's interior. Solar material often streams along these magnetic field lines to produce prominences. Prominences begin and end in the chromosphere, but extend well out into the corona.

Perhaps the most important feature of the chromosphere is the spicule. The chromosphere is covered with spicules. They are vertical streams of chromospheric material moving upward into the corona. The total mass of material moving upward into the corona from spicules is approximately the mass of the entire corona every few minutes. Because the mass of the corona is not rapidly increasing and the solar wind is not that much mass, astronomers know that most of this material must also be falling back into the chromosphere. However, astronomers are not able to observe the falling material. These spicules form a chromospheric bright network, along the boundaries of large supergranules on the Sun's photosphere, that is related to the Sun's magnetic activity.

The Sun's bright magnetic network may play a role in climate changes on Earth. When the Sun is at the maximum of its sunspot cycle and has the largest portion of its surface covered by sunspots, it has a very slightly higher energy output than when it is at sunspot minimum. Even though sunspots tend to reduce the Sun's energy output, brighter areas such as faculae (which become plages when they extend to the chromosphere) and the bright chromospheric magnetic network more than compensate for the energy loss caused by the darker sunspots. During the prolonged Maunder minimum of sunspot activity in the late seventeenth century, Earth's climate experienced its coldest period of the last millennium. From about 1100 to 1250 there was a medieval grand maximum of sunspot activity, corresponding to a warmer than normal period on Earth. Historical climate evidence suggests that solar luminosity variations caused by the Sun's magnetic activity affect Earth's climate. Sunspots and faculae on the photosphere and plages and the bright magnetic spicule network on the photosphere all play a role in changing the Sun's energy output.

KNOWLEDGE GAINED

During a solar eclipse in 1686, not long after the invention of spectroscopy, Jules Janssen, working in France, observed bright emission lines from the chromosphere. One line was very close to the wavelength of a well-known pair of yellow emission lines from the element sodium. This line was so bright that Janssen was able to study it in detail even when there was no eclipse. About the same time, Norman Lockyer, from Britain, also observed this bright line outside eclipses. Further detailed study revealed that this newly discovered spectral line was not quite the right wavelength to be sodium. Furthermore, it did not match the wavelengths of lines from any known element at the time. This newly discovered element was named helium after the word *helios*, for Sun. In 1895, helium was finally isolated on Earth.

As the closest star to us, our Sun is the best studied star and the only star for which it is possible to study surface details. Studying the solar chromosphere as well as the Sun's other regions in detail reveals much about atmospheres of other stars. A more complete understanding of stellar atmospheres, in turn, helps astronomers fine-tune their stellar models and improve their understanding of all aspects of stars.

One of the enduring mysteries about the Sun is why the temperature increases outward in the chromosphere and the corona. The Solar and Heliospheric Observatory (SOHO), launched in 1995, shed light on that question when it observed a magnetic carpet on the Sun's surface. About four thousand magnetic field lines loop up daily from the Sun's interior and then back down into the interior. Because these loops resemble the loops in a carpet, this phenomenon has been called the Sun's magnetic carpet. When these loops burst from the Sun's turbulence, they release energy to heat the outer layers of the Sun's atmosphere.

CONTEXT

Although the chromosphere is a relatively thin layer of the Sun's atmosphere, it is quite important to us. It is not possible to understand the photosphere below the chromosphere, or the corona above it, without understanding the chromosphere. For example, plages in the chromosphere are directly related to faculae in the photosphere. Spicules in the chromosphere extend vertically upward into the lower portions of the corona. Hence photospheric, chromospheric, and coronal phenomena are all interconnected.

The chromosphere plays a role in solar magnetic activity. Interactions between solar magnetic storms and Earth's magnetosphere can affect auroral activity on Earth and, in the case of strong magnetic storms, can affect radio communications on Earth.

The chromosphere and chromospheric phenomena play an important role in the Sun's magnetic activity cycle. If the Sun's luminosity changes with its activity cycle, the chromosphere likely played a role in past climate changes on Earth. It has not been proven, but some scientists think that solar variations may also play a contributing role in Earth's warming. In order to understand this possible role in Earth's climate fully, scientists must have a fuller understanding of the Sun's chromosphere.

Paul A. Heckert

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Chapter 16 of this very readable introductory astronomy textbook covers the Sun.

Frazier, Kendrick. *Our Turbulent Sun*. Englewood Cliffs, N.J.: Prentice-Hall, 1980. A good account of the basic knowledge about the Sun through the publication date.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. Chapter 16 of this introductory astronomy textbook is a complete and up-to-date overview of our knowledge of the Sun.

Golub, Leon, and Jay M. Pasachoff. *Nearest Star: The Surprising Science of Our Sun*. Cambridge, Mass.: Harvard University Press, 2001. A detailed summary of solar science, including a section on the chromosphere that is more detailed than in most.

Heckert, Paul A. "Solar and Heliospheric Observatory." In *USA in Space*. 3d ed. Edited by Russell Tobias and David G. Fisher. Pasadena, Calif.: Salem Press, 2006. Describes the SOHO solar observatory mission, which was used to study the Sun's outer layers, including the chromosphere. This mission revealed much about the Sun's magnetic activity and why the temperature increases in the chromosphere and corona.

Hester, Jeff, et al. *Twenty-First Century Astronomy*. New York: W. W. Norton, 2007. Chapter 13 of this astronomy textbook is about the Sun.

Morrison, David, Sidney Wolf, and Andrew Fraknoi. *Abell's Exploration of the Universe*. 7th ed. Philadelphia: Saunders College Publishing, 1995. The Sun is covered in chapter 26 of this classic astronomy textbook.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. Cambridge, England: Cambridge University Press, 2002. An extremely well written introductory astronomy textbook. Chapter 12 is an overview of the Sun.

Zeilik, Michael, and Stephen A. Gregory. *Introductory Astronomy and Astrophysics*. 4th ed. Fort Worth, Tex.: Saunders College Publishing, 1998. Designed for undergraduate physics or astronomy majors, this textbook goes into more mathematical depth than most introductory astronomy textbooks. Chapter 10 covers the Sun, including the chromosphere.

See also: Coronal Holes and Coronal Mass Ejections; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; Infrared Astronomy; Red Giant Stars; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Corona

Category: The Sun

The solar corona is the outermost, high-temperature portion of the solar atmosphere. This region of the solar atmosphere is the originating site of the solar wind, a flux of charged particles emitted by the Sun, and is considered to be the inner boundary of the interplanetary medium.

OVERVIEW

The solar corona is the outer atmosphere of the Sun. Well known from antiquity, it can be seen with the unaided eye during a total solar eclipse as a glowing white halo around the silhouette of the Moon. The rest of the time, when the Moon is not blocking the bright light from the Sun's photosphere, the corona is too faint at visible wavelengths to be observed without special instruments. The corona was not positively confirmed as a solar feature, rather than an artifact of Earth's atmosphere, until photographs of it were noted to look the same from widely separated sites on Earth.

Spectroscopic studies of the corona show that its visible light emission is a continuous background of all visible wavelengths together with a few superimposed broad emission lines. Thus, its spectrum is quite different from that of the photosphere, the visible solar disk, which has a visible light spectrum consisting of a continuous background with many dark absorption lines, produced by some of the elements found in the Sun. The coronal emission lines did not have wavelengths that matched those identified from any known chemical element, and some astronomers suggested that they were produced by a new element, dubbed "coronium." In 1925, similar lines were observed in the spectrum of a nova (an exploding star), RR Pectoris. In 1939, Walter Grotrian identified a red coronal line as being emitted by iron ions with nine electrons removed (called Fe X, X being the Roman numeral 10), and in 1942, Bengt Edlen identified a green coronal line as being emitted from iron ions with thirteen electrons removed (Fe XIV). It turned out that highly ionized atoms of familiar elements such as iron, calcium, and argon are the source of the coronal emission lines.

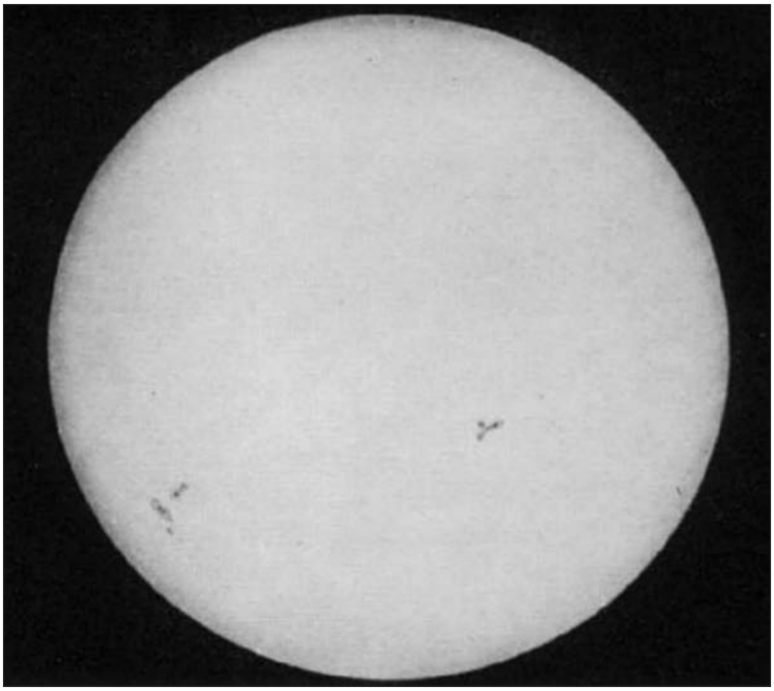
The high degree of ionization requires high temperatures in excess of one million kelvins. It is now thought that this high temperature is produced by the Sun's strong magnetic field. Magnetic waves from the turbulent convective photosphere follow magnetic field lines upward into the lower density corona (about 10^{15} particles per cubic meter or less, compared to 10^{22} particles per cubic meter in the upper photosphere), becoming supersonic shock waves that transfer their energy to the coronal gas. In addition, the complex pattern of magnetic field lines may twist and reconnect, releasing energy. This heating causes the temperature to rise steeply from a minimum of about 4,400 kelvins in the upper part of the photosphere to about 1.5 to 2 million kelvins in the corona. In spite of its high temperature, because of its low density the corona is only about one-millionth as bright as the photosphere at visible wavelengths (which is why the corona usually cannot be seen with the human eye except when the Moon covers the much brighter photosphere during total solar eclipses).

In the early 1930's, a French scientist, Bernard Lyot, perfected the coronagraph, a telescope that artificially created an eclipse of the solar disk. The invention of this type of telescope allowed observation of the corona at times other than total solar eclipses, and for the first time it was possible to follow the evolution of the outer atmosphere of the Sun. The coronagraph is now standard equipment at many ground-based solar observatories, and similar instruments have also been used for studies of the corona performed from orbiting spacecraft. The development of this instrument has been vital to constantly monitoring the corona and studying how it evolves with time. Studies of the corona involving determination of its temperature, density, and chemical composition occupied researchers for the next three decades.

A major advance in solar physics was the publication in 1958 of a classic paper on the dynamics of interplanetary gas, by Eugene Parker of the University of Chicago. He deduced that, as a consequence of the high temperature and low density of the corona, there should be a continual outflow of ionized gas, which he called the solar wind, from its outer edges. Confirmation of the existence of the solar wind was quickly forthcoming from early satellite experiments and, with it, the recognition that Earth is actually surrounded by this flow of charged particles. Earth is exposed to a constant flux of charged particles—mostly electrons, protons (hydrogen nuclei), and alpha particles (helium nuclei)—from the Sun, moving at speeds of a few hundred kilometers per second. However, only a tiny fraction of the Sun's mass is lost via the solar wind, about 10^{-14} of its mass per year.

The corona may be observed not only at visible wavelengths but also in the ultraviolet and X-ray bands of the electromagnetic spectrum. A major advantage of observing at ultraviolet and X-ray wavelengths is that the corona is much brighter in these parts of the spectrum than the photosphere is (just the opposite of the situation at visible wavelengths), so the corona is easily observed against the disk of the Sun. However, Earth's atmosphere effectively blocks ultraviolet and X-ray radiation from reaching Earth's surface, so these wavelength regions must be observed above our atmosphere. Thus it was natural to turn to the brightest source in the sky, the Sun, as the first astronomical object to study with instruments carried above the absorbing atmosphere. The first ultraviolet observations of the Sun were obtained with a spectrograph attached to a fin of a V-2 rocket, which was launched from the White Sands test facility on October 10, 1946. The results demonstrated conclusively the value of sending observational equipment above the absorbing blanket of air and opened a new field of ultraviolet astronomy.

As a step toward the exploration of the solar spectrum in these two wavelength regions, a series of sounding rocket experiments was conducted by the University of Colorado, the Air Force Cambridge Research Laboratory (AFCRL), and the Naval Research Laboratory (NRL) between 1958 and 1965. These experi-



This image of the Sun was made in 1845 as a daguerreotype by physicists Louis Fizeau and Jean-Bernard-Léon Foucault. Sunspots are visible on the surface. (National Science Foundation, High Altitude Observatory)

ments succeeded in characterizing the intensity and the wavelength distribution of the ultraviolet and soft X radiation from the solar corona. Many of these pioneering efforts were undertaken by Richard Tousey of NRL, a leader in this expanding field in the 1960's. The problem with using sounding rockets is that they are above Earth's atmosphere for no more than a few minutes, so long-duration observations are impossible. The solution is to launch spacecraft into orbits around the Earth and in some cases around the Sun, so the Sun may be observed for prolonged periods of time. Some of the first of these were the Orbiting Solar Observatories (OSOs), launched by the National Aeronautics and Space Administration (NASA). Nine of these spin-stabilized satellites were launched over a thirteen-year period, from March, 1962, through June, 1975. The major research gains made through the use of these satellites included a definition of the differences between active solar regions and quiet regions as seen in the ultraviolet, evidence that material found in the corona is cooler over the solar poles than at the equator, and detection of coronal "hole" features, large low-density regions of the corona that typically last for several of the Sun's twenty-seven-day rotation periods. A white-light coronagraph flown on OSO 7 detected the ejection of a huge mass of material upward from the Sun's disk, a mass that evidently had sufficient speed to escape the solar gravitational field.

Initially, NASA intended to conduct a second series of solar investigations from satellites, the Advanced Orbiting Solar Observatory (AOSO) series. These advanced spacecraft were to have improved pointing capabilities and were intended to employ higher spatial resolution for studies of the solar atmosphere. This program was integrated into the crewed Skylab program, and a set of high-resolution solar instruments, including a white-light coronagraph and several ultraviolet and X-ray imaging telescopes, was placed into Earth orbit with the Apollo Telescope Mount (ATM) system on Skylab. Of the many factors that made the ATM-Skylab mission uniquely valuable, two had special bearing on studies of the solar corona. First, the nine-month duration of the mission permitted uninterrupted observation of the evolution and activity of the Sun's atmosphere over a very wide range of wavelengths that are not observable from Earth's surface. Second, since this was a crewed mission, it was possible for the science teams to exploit interactive observing modes, where the data obtained suggested new observations, with the astronauts controlling the instrumentation to optimize the collection of scientific data.

The ATM-Skylab experiments, along with the data and knowledge gained from them about the solar corona, form a cornerstone for modern solar research concerning the outer portions of the Sun's atmosphere. The ultraviolet and X-ray telescopes carried in this orbiting observatory were able to return new images of how the material in the corona is distributed over the disk of the Sun. As never before, astronomers were able to specify the morphology and evolution of the distribution of mass in the outer atmosphere. Coronal holes and their relationship to the solar wind were investigated in detail with ATM data. It also became clear, using the images from the white-light coronagraph, that there is considerable transient activity in the corona, and numerous mass ejection events were detected.

Based on the ATM-Skylab experience, a second white-light coronagraph was launched on the NASA Solar Maximum Mission (SMM) spacecraft in February, 1980. This instrument operated successfully for nine months until it was subject to an electronics failure. The equipment was later repaired in orbit by astronauts who were transported to the satellite by the space shuttle (STS-41C). The electronics package was successfully replaced, and the coronagraph experiment was subsequently operated for a number of years. Thus, the SMM coronagraph has been used to define the variations of the solar corona over a rather long period of time.

KNOWLEDGE GAINED

The physical characteristics of the solar corona are now reasonably well known, as are its basic evolutionary characteristics. The corona is made up of a fully ionized mixture of the solar elements, mostly hydrogen and helium. Metals and other elements constitute a minor fraction of the total mass. The solar mixture is ionized by the high temperature of the outer atmosphere. The high ionization states of the elements identified by their emission lines in the spectrum of the corona require temperatures on the order of 1.5 to 2 million kelvins.

The solar surface exhibits numerous regions in which there is a concentration of magnetic flux. Magnetic forces are frequently strong enough to dominate the motion of the coronal plasma (ionized gas). Unlike Earth's atmosphere, which is controlled by the interaction of pressure and gravitational force, the Sun's outer atmosphere reacts to the interplay of three forces: pressure, gravity, and magnetism. Active regions, associated with sunspots, often show coronal material to be configured into loop or arch patterns, as if the material is confined to specific magnetic field structures. Coronal holes tend to form where the magnetic field lines are open and oriented radially outward into interplanetary space.

The white glow of the corona seen during total solar eclipses is produced by the scattering of photospheric light off the free electrons in the coronal plasma. On average, the electron density near the base of the corona is about 10^{14} to 10^{15} per cubic meter; it is greater near active regions and sunspots and less in coronal hole regions. The major white-light structures tend to be long streamers, which extend from the limb (edge) of the Sun more or less radially, and loops, which are almost always associated with concentrations of surface magnetic fields. Streamers viewed at eclipse are often found to extend outward 4 to 6 solar radii. (The nominal value of the solar radius is 700,000 kilometers, a length approximately equal to twice the distance from Earth to the Moon.) Streamers occur over two kinds of solar disk features: Long, linear streamers occur frequently over magnetically active regions and sunspots, and helmet streamers, which are shaped somewhat like bowling pins, are often seen over magnetic neutral lines, areas in which the magnetic polarity switches from one sign to the other.

The Sun has a twenty-two-year cycle of magnetic activity. Sunspots rise and fall in number with a period which is half of the magnetic cycle, so that concentrations of magnetic flux have been observed to be at maxima in 1958, 1969, and 1980, for example. The total amount of material in the corona is modulated by this cyclic variation of magnetic flux, and the total mass of the corona varies by about a factor of 2 over the sunspot cycle. Also over the sunspot cycle, the number of coronal streamers occurring at any given time varies by a factor of 2.

By observing sunspots and other solar features, it is possible to establish the rotation period of the Sun. Near the equator, it is found to be about twenty-seven days. The basic rotation rate of the corona is approximately equal to the equatorial rotation rate, and the bright features used to follow coronal rotation have lifetimes that range from one to five months. By the late 1930's, it was known that some effect of the Sun dominated Earth's upper atmosphere. Ionospheric disturbances and perturbations of Earth's magnetic field were observed to follow a twenty-seven-day period (like the Sun's rotation). These so-called M-regions could not, however, be correlated with distinct solar structures, such as an enhanced concentration of magnetic field. Following the discovery of the solar wind and its variation in space and time, the problem of locating the solar origins for Earth-perturbing wind streams began to receive much attention from scientists. During the 1970's, the origins of the high-speed solar wind streams were eventually identified with open magnetic field configurations and solar coronal holes, based on the OSO 7 and ATM-Skylab data sets. This correlation constituted a major breakthrough in the task of associating the interplanetary magnetic field and flux configuration at Earth with the physical conditions of the corona at the base of the solar wind.

Along with the solar wind, there are other sporadic ejections of solar material which escape the pull of solar gravity. First detected in the early 1970's, coronal mass ejections (CMEs) were finally explained using data from the white-light coronagraph carried in the ATM-Skylab observatory. Often appearing as huge loops or bubblelike structures having a dimension of half a solar radius in the lower corona, the ejections expand as they rise through the corona and may exceed the size of the Sun at heights above 5 solar radii. The average amount of mass ejected is on the order of 10^{12} kilograms, and their mean outward velocity from the Sun's surface is typically 300 to 400 kilometers per second.

Some CMEs are associated with flare activity, the catastrophic conversion of magnetic energy into thermal energy in or near sunspot regions. Occasionally, CME events occur in areas that have no obvious sunspots. CME events occur about once every three to five days during solar minimum, and the production of events can reach one to two per day during times near the maximum of the activity cycle. CME events constitute one of the most energetic phenomena detected on the Sun; typically, the kinetic energy of such an event can exceed 10^{25} joules.

CONTEXT

The corona provides scientists with a laboratory for studies of how low-density, high-temperature plasmas interact with magnetic fields, and it affords investigators a view of a situation in which pressure, gravity, and electromagnetic forces operate simultaneously. The study of such interactions is known as magnetohydrodynamics (MHD), and the solar corona offers investigators an example at close range where complete MHD processes, such as coronal mass ejections, may be observed over a wide variety of both temporal and spatial scales.

Recognition that Earth is, in fact, subjected to constant bombardment by the solar wind flux has given new impetus to studies of how this wind is generated and how it is controlled by MHD processes. At the center of the solar system, there is a magnetic star (the Sun) that modulates the interplanetary space beyond it. The corona, reflecting the magnetic organization of the Sun over a great variety of spatial scales, is astronomers' best clue to the initial organization of the interplanetary magnetic field and the structure of solar wind flow. Once the locations of coronal hole structures are identified, either by observing in the X-ray or ultraviolet regions or by using limb observations from coronagraphs, it is possible to predict when the sub-Earth point on the Sun is occupied by a coronal hole. This knowledge allows the prediction, with fair accuracy, of when Earth will be subjected to a high-speed stream of solar wind.

The geophysical significance of coronal observations has led to international collaboration in satellite investigations of the Sun. The Solar and Heliospheric Observatory (SOHO) satellite, launched in 1995 and operated cooperatively by the National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA), provides continuous monitoring of the Sun from a stable orbit around the Sun at the Earth-Sun L1 Lagrangian point. (This is the point, located 1.5 million kilometers from Earth along the line between the Earth and Sun, where the gravitational pull of the Sun and Earth are equal.) Among SOHO's many devices for studying various aspects of the Sun, it carries plasma diagnostic instruments for investigation of the solar wind along with several telescopes for solar coronal studies.

Studies of the Sun's corona have provided insights into observed features of other stars. In the late 1970's and early 1980's, X-ray observations from the High-Energy Astronomical Observatories (HEAOs) were interpreted to show, to the surprise of many investigators, that almost all types of stars have outer coronas that radiate at the short wavelengths characteristic of high temperatures. It is now accepted that coronas are a standard feature of stellar atmospheric structure.

Similarly, spectroscopic studies performed in the 1980's with the International Ultraviolet Explorer (IUE) satellite demonstrated that stellar winds, outflows of ionized gas like the solar wind, are common from other stars. However, unlike the Sun, which is a star of moderate luminosity and surface temperature, luminous hot stars have stellar winds driven by the radiation pressure of the intense ultraviolet radiation they emit. Again, insight gained from Earth's own star has aided in the identification and interpretation of a common stellar process.

Richard R. Fisher

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A well-written college-level textbook for introductory astronomy courses. Contains a good description of the solar corona and the SOHO mission.

Eddy, John A. *A New Sun: The Solar Results from Skylab*. NASA SP-402. Washington, D.C.: Government Printing Office, 1979. A summary of the operations of the ATM-Skylab solar observatory. The text is clear and nontechnical, and the book is extensively illustrated with reproductions of many of the most important images returned from this mission.

Foukal, Peter. *Solar Astrophysics*. 2d rev. ed. Weinheim, Germany: Wiley-VCH, 2004. A detailed look at the Sun and our understanding of it. Also covers the history of solar astrophysics. Suitable for undergraduates.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Offers a good description of the solar corona.

Frazier, Kendrick. *Our Turbulent Sun*. Englewood Cliffs, N.J.: Prentice-Hall, 1980. An accessible review of solar physics as of its publication date, this work includes basic discussions of solar-terrestrial interactions and the impact of solar activity on Earth. Illustrations show coronal holes and the interplanetary magnetic configuration. An excellent reference for a reader seeking information on basic problems in solar physics.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook, thorough and well-written. Provides a good description of the solar corona.

Hirsh, Richard F. *Glimpsing an Invisible Universe*. Cambridge, England: Cambridge University Press, 1985. A historical review of the development of X-ray astronomy from the beginnings through the mid-1980's. NASA sources and interviews with experimenters are used as background for this review.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses, divided into short sections on specific topics. Includes a section on the Sun's atmosphere.

Tucker, Wallace, and Riccardo Giacconi. *The X-ray Universe*. Cambridge, Mass.: Harvard University Press, 1985. An interesting history of the development of the discipline of X-ray astronomy by scientists who participated in many of the more significant events of the mid-twentieth century.

See also: Auroras; Coronal Holes and Coronal Mass Ejections; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; Infrared Astronomy; Red Giant Stars; Solar Chromosphere; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Mag-

netic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Evolution

Category: The Sun

The Sun, the closest representative of the stars that populate the universe, has been a reliable benchmark for testing theories of stellar astrophysics for centuries. Much has been learned about the Sun's formation, present status, and future evolution from the developments of modern physics, their powerful theoretical models, and a wealth of observations of both this star and others like it.

OVERVIEW

The Sun was born a little over 4.5 billion years ago from a 10-kelvin, interstellar cloud, roughly 10^{14} kilometers across, of cold atomic and molecular gas. Triggered by an event like a nearby supernova, supersonic turbulence in the cloud caused a compressed region to collapse under its own gravity, fragmenting into smaller (on the order of 10^{12} kilometers) pieces on a timescale of about 2 million years. These clumps flattened into disks from their angular momenta, feeding the disk centers as they collapsed. In the disk that would become our solar system, the center accumulated 1-2 solar masses after an additional 10^4 years and became opaque to its own radiation. Consequently, the central temperature rose to about 10,000 kelvins, and the collapsing mass became a protostar within another 10^5 years.

This protostar, about the size of Mercury's orbit and several times the Sun's current luminosity, contracted further over the next million years. In the process, the early Sun entered the T-Tauri stage of evolution, exhibiting strong protostellar winds and bipolar outflows of jets that became less and less collimated as the surrounding disk flattened and dissipated. These jets' compositions concordantly evolved from being primarily molecular to atomic as the young Sun's temperature rose. This activity subsided as the Sun's protostellar evolution slowed over the next 10^7 years, and gravity struggled to compress the hot, ionized stellar material further. When the core reached a temperature of 10 million kelvins, fusion of hydrogen into helium through the proton-proton chain was possible.

Following a 30-million-year period of slight contraction, the Sun settled into its current state on the main sequence of stellar evolution, with a radius of 6.96×10^5 kilometers, a luminosity of 3.83×10^{36} watts, and surface and central temperatures of 5,780 kelvins and 15×10^6 kelvins, respectively. According to computer models and meteoritic evidence, 4.5 billion years have elapsed since that time. Early in its main sequence history, the Sun increased its luminosity by about 30 percent, as the core temperature and fusion reaction rates rose with the increasing mean atomic weight of the nuclear end products deposited in the interior.

The Sun is expected to remain on the main sequence for another 5.5 billion years, maintaining its normal solar cycles and associated magnetic activity. The Sun's luminosity will also continue to increase slowly as more helium "ash" settles in the core. After a total lifetime of 10 billion years on the main sequence, the Sun's core will be composed of enough helium to shut down hydrogen fusion at its center. Though fusion will still occur in a shell surrounding the depleted core center, hydrodynamic equilibrium will no longer be maintained, and the core will begin to collapse under its own gravity. The subsequent release of gravitational energy will heat the hydrogen-burning shell, increasing the nuclear reaction rates, which in turn will increase the gas pressure on the surrounding solar layers.

Arriving at the subgiant stage in its evolution, the Sun's surface temperature will fall to about 4,000 kelvins, and the solar envelope will expand to a radius three times its current size. The Sun will spend 10^8 years in this stage before becoming a red giant star, 100 solar radii in size—about the size of Mercury's orbit—and several hundred times its present luminosity. The surface temperature will not change appreciably during this time, however.

About 10^5 years afterward, the core will have contracted to the point where the central temperature will reach 100 million kelvins, and helium fusion can occur through the triple-alpha process. The central density of this compressed core will be an extremely high 10^8 kilograms/meter^3, and electron degeneracy pressure will stabilize the core against further collapse. Since a degenerate core is largely insensitive to the additional energy generated by helium fusion, the core will not expand as it is heated, and a runaway "helium flash" will ensue for several hours. The tremendous amount of energy dumped into the core will then heat it to the point where thermal pressure can take over, expanding the core and establishing a new balance with gravity for the next 10^5 years.

The triple alpha process will then proceed at a steady rate in the core, surrounded by a lower-temperature hydrogen-fusing shell. The Sun will reach the horizontal branch of stellar evolution, with a slightly higher surface temperature. The solar envelope will also shrink back to 10 main sequence solar radii. When the core's helium fuel is exhausted, it will be left as carbon ash, surrounded by concentric shells of helium-fusing, then hydrogen-fusing, layers. As before, the core will contract until electron degeneracy pressure dominates over thermal pressure, reaching a central density of 10^8 kilograms/meter^3 and a temperature of 250 million kelvins.

At this red supergiant stage, the central temperature is insufficient to fuse carbon, but the compression will drive the surrounding helium- and hydrogen-burning shells to higher temperatures and luminosities. This will cause the outer layers to expand to 500 solar radii, or about the size of Earth's orbit, cooling to a surface temperature of 4,000 kelvins. The exact size of the Sun at this point is unknown, depending on the severity of mass loss from winds ejected in the red giant phase. The Sun will last a relatively short 10^4 years in this stage. Shell helium burning will happen in a series of violent spurts, causing the solar envelope to fluctuate in size. Additionally, photons produced by electron-nuclei recombinations in the envelope will push the layers out farther with each expansion phase. Eventually, the outer layers will be ejected as a planetary nebula, enriching the surrounding interstellar medium and leaving behind the Sun's compact carbon core. As a white dwarf, this Earth-sized object will cool in a leisurely manner, over many billions of years, to become a black dwarf at a temperature very nearly that of absolute zero.

KNOWLEDGE GAINED

Our understanding of the Sun's history is gleaned mostly from theoretical models. In 1644, René Descartes proposed the theory of vortices, roughly outlining solar genesis from infalling swirling gas. Later, Emanuel Swedenborg's 1734 nebular hypothesis postulated that the Sun was formed by a rotating nebula, an idea further explained in Immanuel Kant and Pierre-Simon Laplace's independently formulated nebular hypotheses. In 1755 and 1796, respectively, they invoked the conservation of angular momentum to picture a collapsing cloud rotating and contracting into a protostellar disk.

In the early twentieth century, James Jeans established the physical criteria governing hydrodynamic equilibrium and the conditions necessary for a cloud to collapse. Current models of star formation favor supersonic turbulence in the parent interstellar cloud, possibly from the shocks of a nearby supernova, as the spark necessary to trigger gravitational collapse. This idea was originally posited by Carl von Weizsäcker in 1944 and Dirk ter Haar in 1950, and it resurfaced in the 1990's with the advent of modern computational power. While still a vaguely understood subject, it is theorized that turbulence was responsible for defining the structure and evolution of the presolar molecular clouds, providing the high compression and transport of angular momentum required for gravity to induce further collapse. The actual isothermal collapse was investigated as a simple case by Richard Larson and Michael Penston in 1969 and also by Frank Shu, who explored the inside-out collapse model that produces protostars. More rigorous investigations of star formation have since been conducted, with thought given to the roles of complex magnetic effects, turbulent viscosities, chemical compositions, and rotation.

Observationally, astronomers have compared these predictions with sunlike stars in various stages of development. Radio observations of the M20 nebula provide images of many stages of stellar evolution, from the parent cloud, fragmentation, and collapse to emission nebulae lit by the first generation of high-mass stars. The 1970's and 1980's saw the discovery of successively lower mass protostars closer and closer to the solar system. For example, observations by the Infrared Astronomical Satellite (IRAS) identified Barnard 5, a currently forming solar-type star. Radio and infrared observations of hydrogen and carbon monoxide have found winds of 100 kilometers per second, as well as expanding knots of water and bipolar radio jets characteristic of protostars.

At higher energies, Chandra, XMM-Newton, and Einstein Observatory X-ray satellites have also observed nascent solar-type stars and star-forming regions for clues to our Sun's past. Although the evolution of the Sun's X-ray luminosity depends on its poorly known initial rotation, astronomers know that it declined gradually from the outset of the main sequence for 100 million years and then dropped by a factor of 1,000 until the present day. This decline is connected to the decline of the solar corona's temperature with time. As for the emerging Sun's immediate environment, it has been suggested that the abundance of neutron-rich iron 60 (Fe^{60}) in some meteorites implies that supernovae were nearby. This would indicate that the Sun was born in a fairly crowded environment similar to the active star-forming regions in the Orion nebula.

The current isolation of the solar system is probably a result of a series of gravitational interactions with other protostars that ultimately ejected the emerging solar system from its crowded neighborhood. Similar isotope-decay analyses have been applied to primordial gas-rich meteorites, deducing the composition and strength of the solar wind within one billion years of the Sun's formation. Observations of current T-Tauri stars in both the X and ultraviolet ranges indicate that the Sun emitted energetic particles and winds as flares at this stage, producing precompacted, irradiated grains with peculiar isotope ratios in the circumstellar disk. Excess neon 21 (Ne^{21}) in meteoritic grains is often seen as evidence for this process. After reaching the main sequence, ancient meteoritic evidence further shows that the solar wind flux gradually declined to its present value. Furthermore, radionuclides in lunar rock samples show that the Sun's proton emission has remained relatively constant over the past five million years, aside from variations from the eleven-year solar cycle. This is also true for heavy ions ejected in flares over this timescale, with the exception that the most ancient flares had an overall enrichment in the trans-iron group of nuclei.

Verification of the Sun's lifetimes in its stages of evolution comes from confirming the computed standard model of solar structure, specifically the nuclear reactions occurring in the core and their observable properties. For instance, the Solar and Heliospheric Observatory (SOHO), launched in 1995, probed the interior composition and temperature structure of the Sun by "listening" to internal pressure waves reflecting off the photosphere. This application of helioseismology has indirectly validated the lifetimes of its various evolutionary stages by substantiating the standard model's predictions of solar composition with depth. Further insight into the Sun's future is gained through observing the evolution of other stars. The Ring and Helix nebulae, for example, are photogenic examples of planetary nebulae ejection and the death throes of Sun-like stars.

CONTEXT

Interestingly, theories of solar formation and evolution seem historically motivated by coincidental developments in physics, taking advantage of the increasing availability and sophistication of quantitative measurements and computations. Early speculations of star formation prior to the nineteenth century, without the advanced physics needed to support them, were easily discarded. In the 1840's, J. Robert Mayer and John James Waterson realized that recently studied chemical and electrical energy sources would be unable to provide the Sun's luminosity for any reasonable timescale. Bolstered by the triumph of thermodynamic principles like energy conservation, William Thomson and Hermann von Helmholtz advocated gravitational contraction as the Sun's energy source. In the twentieth century, solving this problem required the synthesis of two separate fields of science—astronomy and atomic physics—into the new field of astrophysics.

Arthur Eddington and Henry Russell analyzed the interplay of pressure and gravity to understand high-temperature stellar interiors and radiative equilibrium, while Ernest Rutherford and Niels Bohr laid the foundation for solar physics by establishing quantum mechanics. In 1917 and 1920, Eddington and Harlow Shapley concluded that stars must have ages greater than the tens of millions of years allotted by the gravitational collapse scenario.

From 1920 onward, favor shifted from electron-proton annihilation to fusion reactions; these would offer lifetimes measured in trillions and billions of years, respectively. The latter solution won with arguments for a multibillion-year universe from Edwin Hubble's research in receding galaxies, quantum mechanical arguments for the possibility of fusion in hot stellar cores, and Hans Bethe's robust proposal of the proton-proton chain and carbon-nitrogen-oxygen (CNO) cycle (a series of thermonuclear reactions) in 1938. This energy-generation mechanism, along with the associated main sequence lifetime and later evolution, has since been supported by a plethora of increasingly sensitive observational data across the electromagnetic spectrum.

Further progress is expected, especially for uncovering the Sun's early evolution, with the launch and operation of the James E. Webb Space Telescope. This satellite, along with other post-Hubble telescopes, will probe dust-obscured interstellar clouds at infrared wavelengths and at high resolution to help elucidate the intricate picture of low-mass star formation.

Brendan Mullan

FURTHER READING

Green, Simon F., Mark H. Jones, and S. Jocelyn Burnell. *An Introduction to the Sun and Stars*. New York: Cambridge University Press, 2004. A text for introductory-level university astronomy courses or the self-motivated amateur astronomer, discussing the basic physical characteristics of the Sun and other stars. Complex mathematics is avoided.

Lankford, John, ed. *History of Astronomy: An Encyclopedia*. New York: Garland, 1997. A series of very easy-to-digest essays on several key topics in astronomy, within a historical context. Concentrates on the people and social settings behind their science.

Montesinos, Benjamín, Alvaro Giménez, and Edward F. Guinan, eds. *The Evolving Sun and Its Influence on Planetary Environments*. San Francisco: Astronomical Society of the Pacific, 2001. The proceedings of a professional workshop on a range of contemporary topics in solar astronomy, these articles are challenging and technical.

Smith, Michael D. *The Origin of Stars*. London: Imperial College Press, 2004. An excellent text on star formation for undergraduates. The book also doubles as a highly readable and transparent guide for the educated nonspecialist who can loosely follow the meanings of the equations.

Sonett, C. P., M. S. Giampapa, and M. S. Matthews, eds. *The Sun in Time*. Tucson: University of Arizona Press, 1991. This volume is an interdisciplinary collection of essays on the Sun and its impact on the solar system, on a variety of timescales and through many physical mechanisms. Advanced and specialized.

Stix, Michael. *The Sun*. 2d ed. New York: Springer, 2002. A brief introduction to current knowledge of the Sun, meant for the unspecialized scientist-in-training who has a basic conceptual grasp of topics such as thermodynamics and hydrodynamics.

Unsöld, Albrecht, and Bodo Baschek. *The New Cosmos: An Introduction to Astronomy and Astrophysics*. 5th ed. New York: Springer, 2001. A guide to developments in astronomy and astrophysics, geared toward students and researchers. Some prior knowledge of mathematics and physics is recommended.

See also: Auroras; Coronal Holes and Coronal Mass Ejections; Infrared Astronomy; Interplanetary Environment; Nuclear Synthesis in Stars; Red Giant Stars; Solar Chromosphere; Solar

Corona; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Flares

Category: The Sun

A solar flare is a high-energy outburst in the chromosphere of the Sun that emits a variety of electromagnetic radiation ranging from energetic gamma rays to long-wavelength radio waves, along with high-energy charged-particle radiation. If the charged particles reach Earth's atmosphere, they produce various phenomena such as auroras and long-distance communication disruptions.

OVERVIEW

Solar flares are sudden outbursts of electromagnetic and particle radiation in the Sun's chromosphere, releasing from 10^{22} up to 10^{30} joules of energy in a matter of minutes to hours. The intensity of emission rises in a few minutes in catastrophic eruptions, increasing more than ten times in brightness in the visible range alone, while radio, ultraviolet, and X-ray emissions may increase a thousand times. In those few minutes, the brightened area may expand to include a billion square kilometers of the Sun's surface (up to a thousandth of the entire solar disk) with temperatures from ten million up to a hundred million kelvins at the center.

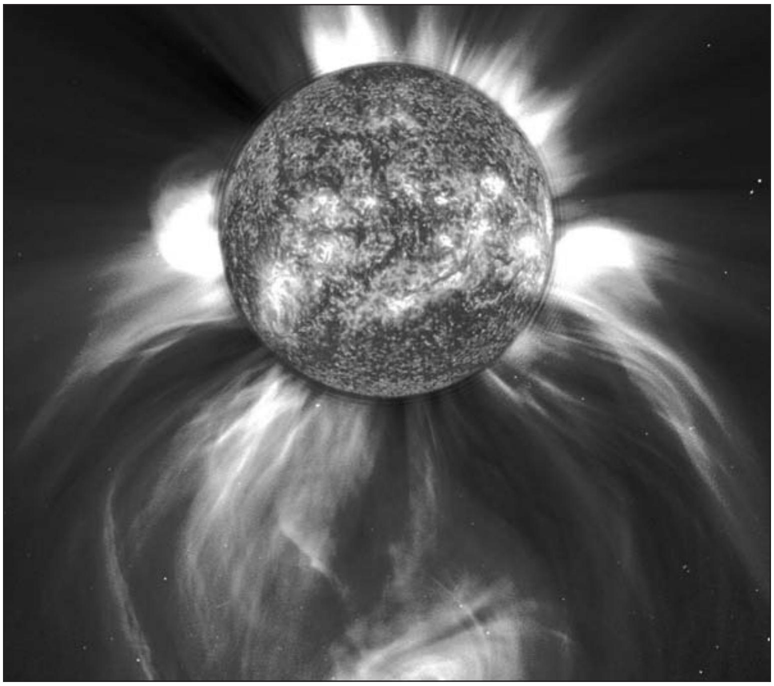
This explosive development usually begins in the upper portion of the chromosphere and then moves upward at a rate up to 100 kilometers per second, often reaching heights of 7,000 to 16,000 kilometers above the photosphere. The brightest flares tend to be the most explosive, reaching a peak in five to ten minutes and then fading over a period of up to two hours.

Solar flares occur where there are strong magnetic fields in the Sun's chromosphere. Large flares are magnetically complex and may have a visible filamented structure. The more magnetically complex the solar environment, the more likely flares are to occur, since colliding and reconnecting magnetic field lines are the most common cause of flares.

The strongest magnetic fields undergoing the greatest changes are associated with and located near the centers of sunspot groups. The intense magnetic fields impede radiation from below the photosphere, creating areas called sunspots, up to tens of thousands of kilometers in diameter, that are cooler and therefore not as bright as most of the photosphere. Sunspot groups usually consist of two main sunspots with opposite magnetic polarities, with many smaller spots clustered around them. Individual sunspots appear and disappear with varying frequency, but their numbers are cyclical, reaching a maximum every eleven years.

Flares occur most frequently above rapidly developing sunspot groups, usually within the first ten to fifteen days of the life of the group. This timing takes place because the most complex magnetic phase of a sunspot group occurs during this period. The flares tend to appear and reappear in association with the same active sunspot regions. Although flares alter the magnetic field, multiple reappearances are possible, which implies that the complex magnetic configuration reestablishes itself between flares.

Another solar feature associated with flares are prominences, glowing plumes or arches of ionized gases that rise above the chromosphere into the corona. They frequently appear as loops of glowing gas that reach from 20,000 to 50,000 kilometers above the Sun's photosphere along magnetic field lines. Other prominences appear as filaments of gas that stream away from sunspots. Prominences, like flares, usually occur in the same latitude belts as sunspots. When they are near spots, they tend to vary rapidly. When away from spots, prominences may be quite stable and last up to three hundred days. Prominences are usually not strong enough to eject matter from the Sun, but eruptive prominences may be so energetic as to propel matter into



Dramatic solar flares and coronal mass ejections in January, 2002. (NASA/ESA/SOHO)

space. Prominences provide some of the charged particles that cause magnetic storms and auroras on Earth, and they are a more modest source of X rays than flares.

Flares emit a wide variety of electromagnetic radiation, generally created by nonthermal mechanisms. Their emissions range from gamma rays and X rays at wavelengths as short as 0.1 nanometer through the spectrum of ultraviolet, visible, and infrared, out to radio wavelengths of up to 10 kilometers. The different kinds of radiation come from different portions of the flare and from different altitudes within it. Thousands of flares occur during each eleven-year sunspot cycle, but flares visible in the white or integrated light portion of the spectrum, are quite rare. The lower in the chromosphere and thus the closer to the photosphere that the flare develops, the more likely that it will be a white-light flare. Only about fifty white-light flares have been observed in the past 150 years.

In addition to emitting electromagnetic radiation, flares emit high-energy charged particles. Magnetic fields in the Sun's atmosphere accelerate charged particles to speeds sufficient to escape from the Sun. Flares eject billions of tons of ionized gas into space at speeds of up to 500 kilometers per second. The ejection of charged particles causes disturbances in the corona called flare surges, which frequently follow the magnetic field lines in the region. Surges increase with the strength of the flare. Flares are one of the sources of shock-wave phenomena that sometimes are observed upon the Sun's surface. At times, the effect of a flare can be traced out to 600,000 kilometers by the effect of the shock wave on thin, gaseous filaments in the Sun's atmosphere. The magnetic field is often significantly altered or dissipated after the flare, although there is a gradual recovery of the magnetic field and the potential for further flares in the same area.

Waves of plasma (charged particles) ejected from solar flares travel outward through interplanetary space. Once they reach Earth, they produce various effects in Earth's atmosphere, such as auroras and long-distance communication disruptions. The fastest particles are high-energy protons moving at an appreciable fraction of the speed of light. A sufficient number of protons may be emitted during a flare to raise the cosmic-ray background on Earth to 180 percent of its normal level.

APPLICATIONS

Scientists are interested in flares because of the fascinating physics associated with their behavior, the grandeur of the size of the phenomena, and the intense dynamism they represent. Their main concern with flares, however, results from the fact that these outbursts can have a significant effect on our planet. Two basic kinds of effects occur: those that are nearly simultaneous with the flare and those that are delayed a day or more.

The simultaneous effects begin at the same time or shortly after the flare is first observed—they are caused by electromagnetic radiation (all forms of which travel at the speed of light) or by high-energy particles traveling almost as fast as light. The most significant simultaneous effects occur in the upper portion of the Earth's atmosphere known as the ionosphere. In this very tenuous gaseous region, the number of free electrons increases, and so the electrical charge increases in an abrupt fashion, called either a "kick" or a "crochet." There is a rapid shortwave radio fade-out that usually begins with the peak of the flare. The fade-out lasts for about twenty-five minutes and then the signal begins to recover. During the fade-out, shortwave radio signals drop to as low as one-tenth of their normal levels. By contrast, the increased reflectivity of the ionosphere for long radio wavelengths of up to 10 kilometers is enhanced. This increased reflectivity also occurs at a lower level in the ionosphere and produces an anomaly in the phasing of the long-wavelength radio waves received directly, compared to those reflected by the ionosphere. Also, the changes in the ionosphere tend to suppress the normal background radio noise.

The delayed effects are caused by the arrival of the particles that travel more slowly and reach the Earth a day or so later. Prominent among the delayed effects are the aurora borealis and aurora australis (the northern and southern lights), which can appear as glowing rays and billowing draperies in the sky. This shimmering display of light is caused when an increased flux of charged particles from the Sun (such as a flare can produce) overloads the Van Allen radiation belts surrounding Earth. The charged particles leak through Earth's magnetic field near the magnetic poles and cascade down into the upper atmosphere. There the high-speed charged particles smash into atmospheric atoms and molecules, exciting them to higher energy states. Then they return to lower energy states by emitting photons of various colors of visible light. Auroras most commonly are seen near Earth's poles, inside the Arctic and Antarctic circles. However, larger flares can produce auroras seen over wider areas; in North America, for example, occasionally they can be seen as far south as the U.S./Mexican border. Not all flares cause auroras, and not all auroras are caused by flares, since the Sun emits charged particles more or less continuously as the solar wind.

Flares send particles in all directions, but Earth receives the most particles when it is directly above the flare—that is, when the flare is near the center of the solar disk as seen from Earth. The most intense magnetic storms will follow such an event.

The magnetic field deflects high-energy charged particles, and the atmosphere blocks high-energy X rays and gamma rays. Together they provide a shield protecting life on Earth. Flares pose a potentially lethal hazard for both space travelers and uncrewed probes that venture beyond the Earth's atmosphere and magnetic field, and better prediction of solar flares as well as better shielding of spacecraft are needed. In addition, the blast of charged particles from a solar flare can increase the density of Earth's thin upper atmosphere, creating extra drag on satellites in orbit. For example, a particularly strong flare on March 6, 1989, caused a drop of a kilometer in the orbit of the Solar Maximum Mission satellite.

The development of ever smaller and denser microcircuitry has made devices using them more susceptible to the effects of high-energy particles. These single-event effects may damage semiconductor circuits. High-energy particles can alter random access memory (RAM) in computers, depositing charge and actually changing stored information and instructions. This has become an increasing problem for computerized control systems, which may eventually be forced to use protective shielding, as is employed with some satellites. A similar but macroscopic problem is that enough charge can be deposited on long metal cables and pipelines to be hazardous, necessitating electrical grounding.

CONTEXT

For a period of five minutes, a little after 11:00 A.M. on September 1, 1859, while mapping sunspots, Richard C. Carrington saw two bright patches of light on the Sun. They moved over the surface of one of the spots he was mapping, with the spot remaining unchanged. Carrington reasoned that he had observed a solar atmospheric phenomenon that had occurred above the sunspot. Independently, R. Hodgson observed the same two spots and also reported them. Their sightings marked the first recorded visual observation of a solar flare in the astronomical literature. Carrington noted that the flare occurred during an intense magnetic storm that lasted from August 28 to September 4, 1859. From then on, flares were occasionally spotted visually. When flares were observed visually through a spectroscope, momentary bright reversals of the dark absorption lines in the solar spectrum were seen. The connection between flares and magnetic storms continued to be noted, and a sense of causation grew among most astronomers, although as late as 1892, the physicist Lord Kelvin disputed the connection. Prior to the 1890's, the distinction between a prominence and a flare was not made.

The systematic study of flares began with the invention of the spectrohelioscope in 1891 by George Ellery Hale. The sunspot cycle peaks of 1936, 1947, and 1958 were studied from ground-based observatories, with the most intense effort made during the International Geophysical Year from 1957 to 1958.

Although much was learned about flares, there was an increasing awareness that there probably were emissions in other parts of the electromagnetic spectrum that could not be detected from Earth's surface because those wavelengths were blocked by the Earth's atmosphere. Instruments borne above the bulk of the Earth's atmosphere by balloons made the first observations of some of these wavelength regions, but the most significant advances came with the advent of the space age. From 1949 to 1960, Herbert Friedman and a team of scientists observed X rays from the Sun through a complete sunspot cycle using sounding rockets to carry detectors above the atmosphere. However, the flights of sounding rockets are suborbital, so the time above the atmosphere was limited to a few minutes.

Much of what is known about flares, their structure, and the mechanisms that cause them has come from satellite-based research, since satellites allow continuous observation. The first satellite devoted to solar studies was Solar Radiation I (1960 Eta 2), launched in August of 1960. The most significant early space-based studies were of X-ray emissions from the Sun. The early results pointed to the areas around sunspots as strong X-ray-emitting regions, with X-ray emission varying dramatically over the eleven-year sunspot cycle.

Another milestone in flare research was the launching of the Solar Maximum Mission (SMM) in 1980. Eight different collectors provided a flow of information from a variety of overlapping wavelengths through the 1991-1992 sunspot peak. The great flare of March 6, 1989, was the largest since satellite observations began. This particular flare released as much energy as 10 trillion 1-megaton hydrogen bombs. The great flare was followed by ten more successive, extremely strong flares. New technology made the thorough study of this series of events possible. Knowledge of the Sun, its features and emissions, continues to expand rapidly with satellite-based observing devices.

A long-standing question about flares has concerned their classification. In the past, some astronomers argued that there was no significant difference among flares; there were simply different power levels of the same basic event. Strong flares allowed observation of all the associated phenomena, while with weaker flares, some phenomena were beyond the ability of existing equipment to detect. Early classification systems simply arranged flares on a scale of one to four, depending on brightness (usually in the visible spectrum). Some scientists added a category of –1 to represent "subflares" (microflares) or exceedingly dim ones. Increased knowledge of the radiation at various wavelengths, especially gamma rays and hard X rays, demonstrated that the brightest visible flares were not the brightest emitters at other wavelengths, and vice versa. This knowledge has led to a new classification system based on several criteria. The letters C, M, and X are used as classes followed by a number from 1 to 15. An X15 flare is the strongest measurable.

Scientists continue to be surprised by the diversity of phenomena that ongoing satellite-based research has revealed. More attention is being devoted to the particle emissions from flares that have an impact on communications.

The expansion of space-based flare-sensing networks will warn of forthcoming communication problems, allowing us to switch to alternative satellite links in order to improve continuity in communications. More attention will also be directed to ultrashort-wavelength electromagnetic radiation, and more study of solar magnetic fields is needed as well.

Ivan L. Zabilka

FURTHER READING

Bai, T., and P. A. Sturrock. "Classification of Solar Flares." *Annual Review of Astronomy and Astrophysics* 27 (September, 1989): 421-467. While highly technical and rather narrowly defined topically, this article remains significant for the more than two hundred bibliographic entries it contains. There is enough description to give the general reader a sense of how complicated the problem of classifying flares remains.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A well-written college-level textbook for introductory astronomy courses containing a good description of solar flares.

Ellison, Mervyn Archdall. *The Sun and Its Influence*. London: Routledge & Kegan Paul, 1955. Primarily of historical interest, since much of the book is outdated, chapter 5 discusses what was known of the physics of flares in a narrative form.

Emslie, A. Gordon. "Explosions in the Solar Atmosphere." *Astronomy* 15 (November, 1987): 18-23. An excellent popular survey of the kinds of research under way in the late 1980's. Provides clearly written background information on the nature of flares themselves.

Foukal, Peter. *Solar Astrophysics*. 2d rev. ed. Weinheim, Germany: Wiley-VCH, 2004. A detailed look at the Sun and our understanding of it. Also covers the history of solar astrophysics. Suitable for undergraduates.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses, containing a good description of solar flares.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook with a good description of solar flares.

Kundu, M. R., B. Woodgate, and E. J. Schmahl, eds. *Energetic Phenomena of the Sun*. Boston: Kluwer, 1989. A volume in the Astrophysics and Space Science Library that was the result of three meetings held at the Goddard Space Flight Center in January and June of 1983 and February of 1984. While highly technical, it offers some descriptive passages of interest to the general reader. Especially valuable for its bibliography.

Maxwell, Alan. "Solar Flares and Shock Waves." *Sky and Telescope* 66 (October, 1983): 285-288. Includes a summary of information about flare-generated shock waves as a source of acceleration for particles emitted by the Sun. Considers some difficult concepts and presents them clearly.

Meadows, A. J. *Early Solar Physics*. Elmsford, N.Y.: Pergamon Press, 1970. A brief historical account containing a fine summary and significant information about related topics. The original reports of Carrington and Hodgson are reprinted in appendixes.

Rust, David M. "Solar Flares, Proton Showers, and the Space Shuttle." *Science* 216 (May 28, 1982): 939-946. Surveys the dangers of radiation from flares to occupants of the space shuttle, suggesting that no adequate level of shielding is possible. Excellent diagrams; only semitechnical.

Ryan, James M. "The Solar Maximum Mission." *Astronomy* 9 (May, 1981): 6-16. An excellent overview of the state of knowledge about solar flares in 1980. Surveys the results from the Solar Maximum Mission during the ten months before the spacecraft failed. Filled with readable information.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A thorough college textbook for introductory astronomy courses, with a section on solar activity and flares.

Verschuur, Gerrit. "The Day the Sun Cut Loose." *Astronomy* 17 (August, 1989): 48-51. An excellent popular summary of the magni-

tude and effects of the great flare of March 6, 1989. Contains a reliable summary of information about the general nature of flares.

See also: Auroras; Coronal Holes and Coronal Mass Ejections; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth-Sun Relations; Interplanetary Environment; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Van Allen Radiation Belts.

Solar Geodesy

Category: The Sun

Solar geodesy is the study of the size and shape of the Sun. In attempts to measure the shape of the Sun precisely, astronomers accidentally discovered solar oscillations, complex rhythmic pulsations involving both the deep interior and the atmosphere of the Sun.

OVERVIEW

Geodesy is the mathematical study of the size and shape of the Earth, and how these affect the precise location of points on the Earth's surface. Solar geodesy is the application of this discipline to the Sun, especially the study of the size and shape of the Sun. The main impetus for such studies up to the mid-1900's was twofold: (1) to see if the Sun is oblate due to its rotation, and (2) to find out if the Sun is slowly shrinking. Although not conclusively settled, it appears neither of these actually is the case. However, detailed measurements revealed something unexpected: The Sun undergoes complex oscillations, or as one researcher put it, "The Sun rings like a bell." This discovery led to a new branch of solar studies called solar seismology, or helioseismology, which uses these oscillations as probes of the Sun's interior, analogous to the way geologists use seismic waves as probes of the Earth's interior. The Sun's vibrations are geometrically complex, and solar physicists can infer the physical nature of the solar interior by analyzing the timing and amplitudes of the many vibration patterns observed on the Sun's surface. Hence, solar geodesy today is focused largely on helioseismology.

In 1960, Robert Leighton of the California Institute of Technology observed that small regions on the Sun's surface were oscillating or pulsating with a period of approximately five minutes. He detected this oscillation by using the Doppler effect, a shifting of the wavelengths or frequencies of electromagnetic radiation. Motion of the source toward the observer causes a "blueshift," a shift to shorter wavelengths. Motion of the source away from the observer causes a "redshift," a shift to longer wavelengths. The amount of the shift is in proportion to the speed of the source toward or away from the observer. Leighton wondered whether the pulsations he observed were occurring only in small regions on the Sun, perhaps a few thousand kilometers in extent, or over a more extended region, possibly even the entire Sun. Although his discovery was noted and subsequently confirmed by many other observers at many observatories, these confirming observations could indicate only that the oscillations were localized. The data would not permit any conclusion regarding coherent motion (connected or related motion over a wide region) on the Sun.

In the early 1960's, a group of researchers at Princeton University led by Robert H. Dicke claimed that they had measured an oblateness (a distortion of a sphere resulting from compression along the polar axis and stretching around the equator) in the shape of the Sun. Using a highly specialized telescope of their own construction, they made measurements of the solar equatorial and polar diameters. They claimed the equatorial diameter to be slightly greater than the polar diameter.

They made these measurements to test Albert Einstein's general theory of relativity against an alternative, the Brans-Dicke scalar-tensor theory, developed by Dicke and Carl Brans. Both theories predicted, among other things, that the orbit of the planet Mercury around the Sun should precess, meaning that Mercury's elliptical orbit itself should slowly rotate around the focus occupied by the Sun. According to the theory of general relativity, this would be due to the warpage of space-time by the Sun's mass. According to the Brans-Dicke theory, however, a small oblateness in the shape of the Sun would cause the same thing. Thus, the oblateness measurements were crucial in distinguishing between general relativity and the Brans-Dicke scalar-tensor theory.

In the mid-1960's, Henry Hill at the University of Arizona designed and built another specialized telescope for detecting distortions in the Sun's surface. Hill's preliminary measurements, however, could not confirm the measurements of the Princeton group. By the late 1960's, other astronomers had shown that the Princeton results probably were due to solar activity that was producing increases in brightness in the Sun's equatorial regions. The increased brightness was caused by plages—bright, patchy regions on the Sun produced by magnetic activity. The plages are usually found near sunspots or centers of magnetic activity, which tend to concentrate within about 30° to 40° north and south of the Sun's equator. Measurements of the solar equatorial diameter are highly influenced by plages. There is a tendency to overestimate the edge of the Sun, or its limb, because of the bright glow of these plages near the solar equator.

In 1975, Hill and Robin Stebbins concluded that the Sun was not oblate but apparently fluctuated or oscillated rhythmically over a large region, possibly its entire surface. Further research verified that the observed phenomena were genuinely solar, not introduced by Earth's atmosphere or telescopic effects. Delicate instruments always have random fluctuations, or "noise," associated with their measurements, but researchers showed convincingly that the observed oscillations were real and not simply the misinterpretation of observational noise.

It came to be accepted that the Sun shakes or vibrates in a range of spectacular ways like a ringing bell. In this process, the Sun's shape undergoes tiny, patterned distortions of a rhythmic nature. These distortions are in effect three-dimensional waves that pass from the deep interior of the Sun to the surface and also move about the Sun's circumference. The periods of these measured oscillations range from about 3 minutes to 160 minutes and perhaps longer.

These oscillation patterns carry information about the deep solar interior, which cannot be observed directly since the Sun's interior is opaque. The waves or disturbances producing these oscillations originate at different levels within the Sun, some just beneath the surface and others farther within the interior. The observed properties of the oscillations—their timing and the extent of their displacement—depend upon the environment through which they pass on their way to the surface, where solar physicists detect them spectroscopically using the Doppler effect. Just as geophysicists study seismic waves traveling through Earth as a result of earthquakes, solar physicists study solar oscillations traveling through the Sun in order to study the solar interior.

Why does the Sun shake? What in the solar interior sets the oscillations in motion, and how do they move through the Sun? All physical bodies, whether they be solid or fluid (fluids include gases and liquids), can oscillate or shake with a variety of frequencies or periods. Virtually any disturbance or natural internal motion can start the oscillations. Convection (the process whereby heat is transferred in the outer regions of the solar interior) is one such stimulus. Solar convection involves the ascent of hot, lower-density bubbles of gas. These bubbles are heated in the deeper, hotter interior; with lower density, they rise buoyantly upward and convey the heat to higher, cooler layers. The motion of the bubbles disturbs the surrounding gases and starts them oscillating, and the oscillations move throughout the Sun. Just as the length of an organ pipe determines the note played, certain fractions of that length produce overtones or harmonics. The Sun's spherical shape, interior density, and temperature determine both the frequency and the length of the waves that are set in motion, regardless of the process that caused those waves.

In a structure such as the Sun, the oscillations can be of two types, depending on the nature of the force that maintains the oscillation. Either gravity or pressure can supply the restoring force (the mechanism that brings the displaced fluid back to its original position, thus maintaining the motion necessary to produce the pulsations or waves). Small pressure fluctuations give rise to acoustic or soundlike waves, alternate compressions and rarefactions that move through the fluid at the speed of sound. Waves of this type are referred to as "p modes." Just below the visible solar surface, the convection produces a deafening roar, similar to that produced by a jet or rocket engine.

Gravity waves are created when an element or small volume of fluid is displaced and subsequently returned to its position by gravity. This type of wave can occur only when the density or compactness of the material varies with depth. On Earth, water waves larger than small ripples are of this type. The wave moves toward the shore, and its vertical displacement is restored by gravity, causing the wave to move along the surface of the water. (These gravity waves are not to be confused with the gravitational waves predicted by Einstein's theory of relativity, which are of a completely different nature.)

Any movement or perturbation in the interior of the Sun can start the quivering process and produce an oscillation. Individual oscillations can be manifested in many ways, with a variety of wavelengths and nodes (regions or points free of oscillations; in a vibrating string, the node would be the tied-down point not undergoing vibration). Because of the three-dimensional, spherical structure of the Sun, a virtually endless variety of modes or patterns can occur in it. Some modes encompass the Sun's entire structure, while others take place in localized regions.

The quivering waves generated by these oscillations reflect from the surface layers of the Sun and speed back toward the deep interior, where in turn they are refracted or bent back toward the surface. The layer of refraction depends on the speed of the waves. In many cases, the speed is the local speed of sound for the solar interior gas, which in turn depends on the temperature of the gas at that location. The reflection from the surface is caused by a rapid decrease in density, since the surface represents an abrupt boundary between the solar gas and space. Since sound cannot travel through a vacuum, the waves are reflected at the boundary back into the Sun. Such a process can occur again and again, resulting in waves moving along curved paths and bouncing and refracting their way completely around the interior of the Sun.

KNOWLEDGE GAINED

The Sun oscillates like a giant, spherical bell. This discovery was first made by Robert Leighton in 1960. The observed oscillations were about five minutes in duration. It was not known at that time whether the oscillatory motion involved small, localized regions on the Sun or much larger regions. Hill's 1975 observations found oscillations of relatively short duration, with periods of a few minutes. A considerable effort went into showing that these oscillations were real and not the result of distortions caused by Earth's atmosphere or by instrumental effects. By 1980, however, similar oscillations had been observed at other observatories, and some had been shown to have periods between two and three hours long.

These oscillations were found to be of two types: (1) pressure waves and (2) gravity waves, based on the natural force that maintains the oscillations. In the case of short-period waves, those lasting a few to several minutes, the wave is essentially a sound wave traveling through the solar interior. Such waves are known as pressure waves, since alternate compression and rarefaction of local gas produces the oscillations. Long-period waves, with oscillation periods as long as two or three hours, have gravity as the driving, restoring force. A single oscillation can be set up in an almost infinite number of geometric shapes, or modes. Many of these patterns have striking geometrical beauty and symmetry.

Observations of solar oscillations have revealed many new things about the nature of the Sun's interior and have confirmed other things. For example, at the surface, the Sun's equatorial zone is observed to rotate somewhat faster than do regions north or south of the equator; this phenomenon is known as differential rotation. The rotation of the Sun's interior cannot be directly observed, but observations of surface waves generated in the deep and shallow solar interior indicate a complex rotation pattern, with "flowing rivers" and "zones" and layers rotating at different and sometimes varying rates. Other studies find oscillation patterns that are consistent with the Sun's deep interior temperature being about 15 million kelvins, a value derived from computer models of the Sun's interior.

Early theories of stellar evolution and energy generation held that stars should slowly contract gravitationally as they age, so solar geodesy once was concerned especially with measuring such shrinkage in the Sun. Solar eclipses provided much of the raw data. If the Sun were shrinking, annular eclipses would have been more common and the duration of totality during total eclipses would have been shorter in the past. Although some early studies claimed to have found these effects, they have not been substantiated by more recent work that better accounts for instrumental errors.

CONTEXT

Later studies concentrated on measuring solar oblateness, but claims that the Sun is slightly oblate have not been confirmed. Instead, solar observers have discovered that the Sun oscillates in complex patterns. By observing oscillations visible on the Sun's surface, astronomers are assembling a detailed picture of the solar interior, in much the same way that geologists and geophysicists gain knowledge of Earth's interior by the study of waves produced in earthquakes. The most extensive series of observations of solar oscillations have been made by the Global Oscillations Network Group (GONG), an array of small telescopes around the world that provide continuous monitoring of solar oscillations over many years. The data gathered by GONG are enabling astronomers to differentiate real solar oscillations (together with their amplitudes and periods) from noise produced by instrumentation or Earth's rotation. An average value for the diameter of the Sun is 1,392,000 kilometers, or the equivalent of 110 Earths lined up side by side.

James C. LoPresto

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Good description of solar oscillations.

Foukal, Peter. *Solar Astrophysics*. 2d rev. ed. Weinheim, Germany: Wiley-VCH, 2004. A detailed look at the Sun and our understanding of it. Also covers the history of solar astrophysics. Suitable for undergraduates.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Includes a good description of the Sun's atmosphere and energy output.

Frazier, Kendrick. *Our Turbulent Sun*. Englewood Cliffs, N.J.: Prentice-Hall, 1980. This book summarizes the research and observational discoveries made about the Sun from about 1950 to 1981. When possible, the author relates these developments to the impact the various phenomena have on Earth, considering questions such as whether the solar activity cycle affects Earth's climate.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook with a good description of solar oscillations.

Giovanelli, Ronald. *Secrets of the Sun*. Cambridge, England: Cambridge University Press, 1984. This book outlines modern solar physics with the aid of spectacular black-and-white images from numerous observatories. Giovanelli focuses on the physical nature of the various atmospheric components of the Sun: the photosphere, the chromosphere, and the corona.

Harvey, John W., James R. Kennedy, and John W. Leibacher. "GONG: To See Inside Our Sun." *Sky and Telescope* 74 (November, 1987): 470-476. This article examines the field of helioseismology, using detailed drawings and diagrams. The article outlines an ambitious project to put into place small automated telescopes around the world to study solar oscillations over a five-to-ten-year period.

Leibacher, John W., Robert W. Noyes, Juri Toomre, and Roger K. Ulrich. "Helioseismology." *Scientific American* 253 (September, 1985): 48-57. Colorful diagrams illustrate a detailed description of how acoustic waves within the Sun cause the surface to heave up and down. The waves' timing and geometrical patterns are described, with emphasis on how study of them leads to knowledge of the structure, composition, and dynamics of the invisible interior of the Sun. Somewhat challenging reading.

See also: Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Infrared Emissions

Category: The Sun

Most solar infrared and far-infrared radiation is emitted from the coolest layers in the solar atmosphere, which are found in the upper portion of the photosphere. Analysis of this radiation not only allows scientists to understand these important layers of the solar atmosphere but also provides observational confirmation of the simplest known interaction between matter and radiation: local thermodynamic equilibrium.

OVERVIEW

Infrared (or IR) radiation is a form of electromagnetic radiation between visible light and microwaves (short-wavelength radio waves). Gases in Earth's atmosphere (especially water vapor) absorb much of the IR radiation—particularly the "far infrared," the longer-wavelength IR—before it reaches the ground. Effective study of IR radiation from astronomical sources must be done using telescopes at dry, high-altitude, mountaintop observatories above the densest part of Earth's atmosphere, or preferably using space telescopes completely outside our atmosphere.

The IR radiation emitted by the Sun represents only a tiny fraction of the total solar radiative energy. The rate at which the Sun emits IR energy is only about 0.058 percent of the total solar luminosity, which is the rate at which the Sun radiates energy over the entire electromagnetic spectrum. The solar constant is the average total solar irradiance (electromagnetic energy per unit time per unit area) impinging on the top of Earth's atmosphere. Its value is 1,368 ± 7 watts per square meter, as measured by the Solar Maximum Mission (SMM). Of this, only 0.802 ± 0.026 watts per square meter is in the IR.

The IR is the simplest part of the solar spectrum. The radiation is almost continuous over the entire IR spectral range, with very few dark absorption lines. The continuous background spectrum is thermal radiation, such as a blackbody (a perfect thermal radiator) emits. Absorption lines (called Fraunhofer lines in the solar spectrum) are formed when photons with specific energies are absorbed by electrons jumping to higher energy levels in atoms or molecules, or by molecules going to higher-energy rotational or vibrational states. Consequently, less energy remains in the spectrum at the wavelengths corresponding to the absorbed photon energies, so the spectrum looks darker at those wavelengths. The absorption lines that are present serve as a chemical fingerprint revealing that specific elements and compounds are present in the source. However, the absence of absorption lines in some part of the spectrum—for example, in the IR—does not necessarily indicate certain elements or compounds are not present, just that the physical conditions are not right for them to absorb in that spectral range if they are present in the source.

IR radiation arises primarily in the upper photosphere (the visible surface of the Sun) and the lower chromosphere (the layer of the Sun's atmosphere, a few thousand kilometers thick, immediately above the photosphere). These levels of the solar atmosphere consist of homogeneous strata. From place to place and layer to layer, the gas and its behavior show a remarkable similarity. This is where the solar atmospheric temperature falls to its minimum value and begins to rise to the higher temperatures found in the chromosphere and corona, the extensive outer layers of solar atmosphere above the photosphere.

The photosphere consists of a series of layers from which most of the Sun's visible light is emitted into space. These layers make up a very thin shell, only a few hundred kilometers thick. The opacity of the photosphere increases rapidly with depth; thus the intensity of emitted radiation drops off rapidly with depth into the photosphere. A photon emitted outward from the lower photosphere has a large probability of being absorbed or scattered by the atoms and free electrons within the photosphere. Because of this, photons are likely to escape into space only from the photosphere's uppermost layers. For this reason, the edge (or limb) of the Sun is sharply defined, and the Sun appears to have a definite surface.

The photosphere produces the continuous radiation observed across the entire solar spectrum. The intensity peaks in the yellow-green part of the visible spectrum and falls off toward both longer and shorter wavelengths. That is why the IR contributes such a small percentage of the Sun's total electromagnetic radiation. This spectral distribution of intensity is similar to that of a blackbody (an ideal thermal radiator) with a temperature of about 5,800 kelvins. In reality, each layer making up the photosphere emits its own blackbody spectrum, which in turn depends on the temperature of that layer. The sum total of all the emissions from all the layers is similar to one imaginary layer emitting at approximately 5,800 kelvins; this is referred to as the Sun's "effective temperature." Those layers that are deeper in the photosphere emit at higher temperatures, and those that are higher in the photosphere emit at lower temperatures, as low as 4,400 kelvins.

The blackbody spectral distribution results from the high opacity of the layers. (Perfect blackbody radiators are perfectly opaque.) The source of this high opacity is the presence of negative hydrogen ions, first proposed by Rupert Wilt in the late 1940's. The negative hydrogen ion is a hydrogen atom with an additional (second) electron weakly attached. It easily absorbs radiation in the visible spectrum and especially in the infrared. The second electron is bound to the hydrogen atom very weakly—with a bond about 3.5 percent as strong as that of the first electron. Since the bond of the second electron is so weak and since the temperature drops off rapidly toward the top of the photosphere, the electron density also diminishes rapidly in the photosphere's upper reaches. In other words, the number of free electrons and negative hydrogen ions per unit volume—known as their "number densities"—are very sensitive to temperature and thus height in the photosphere.

The phenomenon of limb darkening can be observed on any good white-light solar image (an image of the Sun obtained over the entire range of visible wavelengths, the combined colors from violet to red giving a white image with a slight yellow tinge). The limb of the Sun is noticeably darker than the center of the solar disk because of the temperature gradient in the photosphere. As an observer's line of sight moves toward the limb, it passes through only the upper, cooler layers, whereas deeper, hotter layers are seen near the disk center. Thus the intensity of light decreases toward the limb due to the drop in temperature with height. (In contrast, the Sun's limb appears brighter than the disk center in observations of the chromosphere and lower corona, because of the reversal of the temperature gradient in those layers.)

Since intensity of radiation can be measured as a function of both distance from the disk center and wavelength, much can be learned about the way in which physical properties change with depth, once the processes of photon emission, absorption, and scattering within the atmosphere are understood. These processes control the flow of electromagnetic radiation through the solar atmosphere by radiative transfer. Using the processes of radiative transfer and observed details of the solar spectrum and limb darkening, astronomers can calculate mathematical models of the solar atmosphere. These models are tabulations of height versus variables of interest such as temperature, density, and pressure.

KNOWLEDGE GAINED

The entire solar photosphere emits radiation like a stacked system of glowing shells, each shining at its characteristic temperature. The net effect of all these glowing shells is similar to that of one thin shell emitting at a temperature of 5,800 kelvins, the "effective temperature" of the photosphere. The infrared and far-infrared spectra are emitted from the coolest layers (4,400 kelvins), located at the top of the photosphere. Most of the solar IR spectrum is in a continuous form, with few dark absorption lines. IR radiation originates mainly in the uppermost photosphere. At this level, the solar atmosphere is at its coolest temperature: about 4,400 kelvins.

CONTEXT

The study of the IR spectrum is important because these observations, along with limb-darkening measurements, are needed in modeling the solar atmosphere. These layers of gas are essentially in local thermodynamic equilibrium (LTE). Radiation is transferred in LTE by well-understood processes: photon emission, absorption, and scattering. The primary contributor to the opacity of all layers of the photosphere is the negative hydrogen ion. Sources of opacity and transport of energy by known nonthermal processes—such as mechanical mechanisms, sound waves, shock waves, and magnetic energy—are very small in the photosphere.

Furthermore, spectral studies of solar IR have illuminated details of the atomic and molecular composition of Earth's atmosphere. These studies, in turn, help scientists to understand the interaction of other bands of electromagnetic radiation, such as ultraviolet, with the terrestrial atmospheric constituents detected in the IR.

James C. LoPresto

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. This well-written introductory astronomy textbook contains a good description of the Sun's atmosphere and energy output across the electromagnetic spectrum.

Foukal, Peter. *Solar Astrophysics*. 2d rev. ed. Weinheim, Germany: Wiley-VCH, 2004. A detailed look at the Sun and our understanding of it. Also covers the history of solar astrophysics. Suitable for undergraduates.

Frazier, Kendrick. *Our Turbulent Sun*. Englewood Cliffs, N.J.: Prentice-Hall, 1980. This book summarizes the research on and observational discoveries about the Sun from about 1950 to 1981. When possible, Frazier relates these developments to the impact on Earth of the various phenomena, such as whether the solar activity cycle affects Earth's climate. Suitable for general audiences.

Giovanelli, Ronald. *Secrets of the Sun*. Cambridge, England: Cambridge University Press, 1984. Spectacular black-and-white images from a number of observatories accompany an account of modern solar physics. Giovanelli highlights the physical nature of the various atmospheric components of the Sun: the photosphere, the chromosphere, and the corona.

Nicolson, Iain. *The Sun*. New York: Rand McNally, 1982. In an atlas format, Nicolson provides a historical perspective on solar studies; discusses the relationship of the Sun to the stars, the Galaxy, and the universe; and presents a detailed description of the solar interior and atmosphere. Solar activity and solar-terrestrial relationships are examined at length. For readers with a basic background in solar physics.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses. Has several sections on the Sun's atmosphere and energy output.

See also: Coronal Holes and Coronal Mass Ejections; Earth-Sun Relations; Electromagnetic Radiation: Thermal Emissions; Infrared Astronomy; Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radi-

ation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Interior

Category: The Sun

The nuclear fusion reactions that power our Sun occur in its interior. The nature of this interior determines conditions at the surface of the Sun, affecting the rest of the solar system.

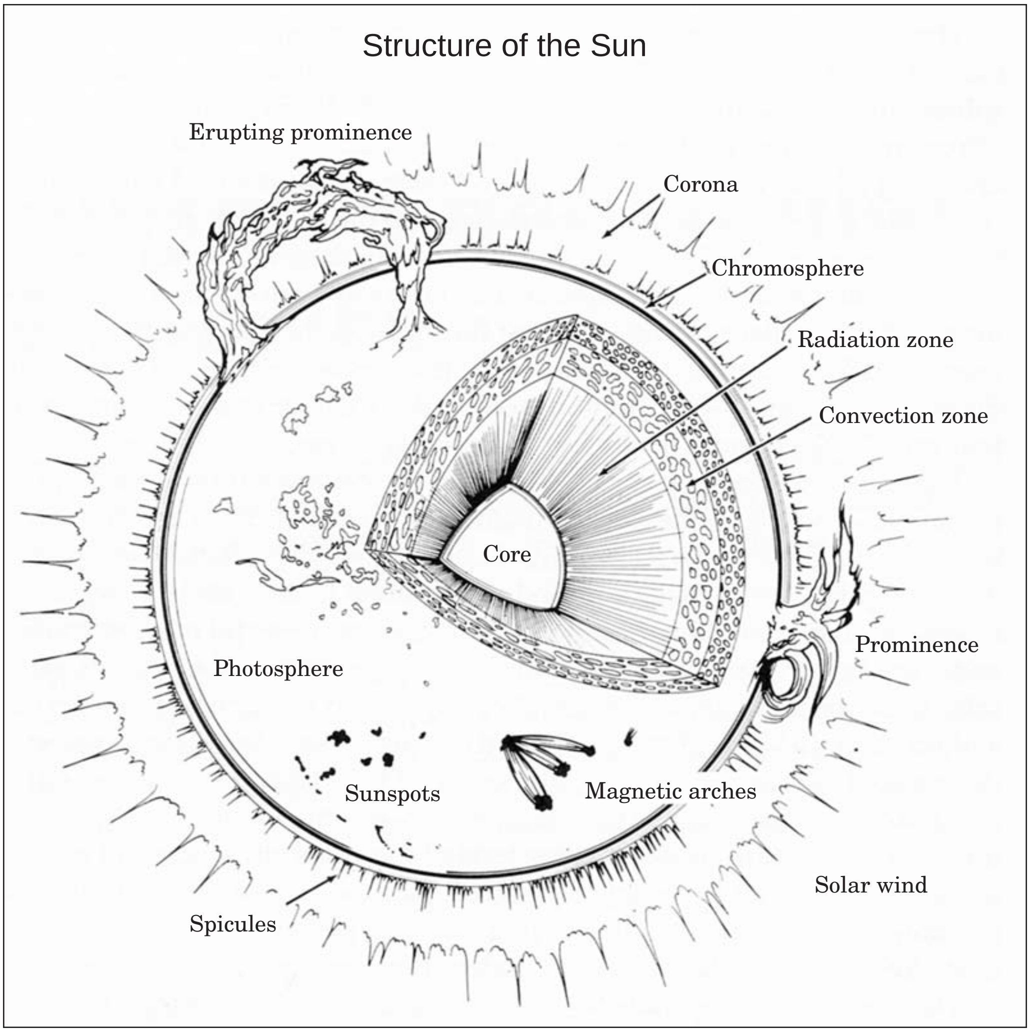
OVERVIEW

The Sun is composed primarily of hydrogen and helium. Though these substances are gaseous on Earth, the conditions in the Sun are such that they do not behave inside the Sun the way that they do on Earth. The Sun's gravity compresses and heats these gases in the Sun's interior. About one-fifth of the way below the surface of the Sun, the gases are so hot that they are ionized—that is, they have been stripped of their electrons. The closer to the center of the Sun these gases are, the more compressed and heated they are. Almost 94 percent of the Sun's mass is contained within the innermost half of the Sun's radius.

Near the center of the Sun, the temperature is more than 15 million kelvins. Gases are compressed to a density of about 160 grams per cubic centimeter (g/cm^3), more than fourteen times denser than lead. At such density and temperature, hydrogen nuclei begin to fuse into helium. Every second, close to 600 million tons of hydrogen is consumed in this fusion process. The result of this fusion is the production of about 596 million tons of helium. The difference in these two mass figures, 4 million tons, is converted into energy via the equivalence of mass and energy given by Albert Einstein's famous equation $E = mc^2$. This energy released in the deep interior of the Sun is what supports the Sun against gravity and keeps it from collapsing further under its own weight.

Density and temperature are greatest at the center of the Sun. This is where nuclear fusion occurs at the fastest rate. The density and temperature decrease with distance from the center of the Sun. Therefore, the rate of nuclear fusion decreases with increasing distance from the center of the Sun. Beyond about 25 percent of the distance from the Sun's center to its surface, the density and temperature are too low to support fusion at any appreciable rate. This innermost portion of the Sun, where nuclear fusion occurs, is called the core of the Sun. The temperature at the top of the core is about 7 million kelvins. However, there is no well defined edge to the core. Rather, the farther from the center of the Sun, the lower the rate of fusion.

Throughout the core, energy is produced through nuclear fusion. Much of this energy is initially in the form of high-energy gamma rays. The gamma rays produced in this manner travel only an extremely short distance before they collide with an electron. Gamma rays are a form of electromagnetic radiation. While often described as waves, electromagnetic radiation also acts like particles, called photons. These photons carry momentum. Therefore, when the gamma rays collide with electrons, they scatter off of the electrons in a process called Compton scattering. In the scattering process, the gamma rays lose momentum and energy and the electrons gain momentum and energy. These collisions between the gamma rays and the electrons are what support the interior of the Sun. Gamma rays continually rebound from electron to electron, scattering in random directions. Gamma rays continually lose energy, eventually becoming X rays. The radiation scatters in random directions with each collision. In this manner, called radiative diffusion, the energy from the nuclear fusion in the Sun's core gradually works its way outward from the middle of the Sun. Because of the large number of collisions in random directions, it takes a long time for the radiation to travel outward. On average, the radiation diffuses outward at a rate of about 50 centimeters per hour and takes nearly 170,000 years to make its way out of the Sun.



Source: Bevan M. French and Stephen P. Maran, eds. *A Meeting with the Universe: Science Discoveries from the Space Program*. NASA EP-177. Washington D.C.: National Aeronautics and Space Administration, 1981, pp. 68-69.

Radiative diffusion dominates until a distance from the center of about 71 percent of the Sun's radius. At that point, the temperature has dropped to about 2 million kelvins and the gas density has dropped to about 0.2 g/cm^3 (about 150 times denser than Earth's sea-level atmospheric density). Here some electrons are captured by the atoms. Instead of the light scattering off of electrons, it is absorbed by atoms, heating the gases. The hot gas then expands and rises. When the gas reaches the surface of the Sun, the photosphere, it cools by radiating light into space. The cooler gas then sinks until it is again warmed by absorption of more energy. Heat transfer in this manner is called convection. Since convection dominates the energy transfer mechanism in the upper portion of the Sun, the top 29 percent of the Sun's radius is called the convective region. The lower 71 percent of the Sun's radius, including the core, is called the radiative region because radiative diffusion is the chief mechanism for energy transfer.

Between the radiative region and the convective region is a small region that acts as an interface between the radiative and convective regions. This zone is called the tachocline. Below the tachocline, atoms are almost all ionized, and radiative diffusion dominates. Above the tachocline, atoms have electrons, and convection dominates, but the transition is not sharp. There is no set distance from the center of the Sun where the transition between radiative diffusion and convection occurs. The lower portion of the tachocline has more radiative diffusion than convection, and the upper portion of the tachocline has more convection than radiative diffusion. The tachocline is important to solar astrophysicists because this is the region where the Sun's magnetic field is believed to be produced.

The top of the convective region is the photosphere. This is often regarded as the visible surface of the Sun. However, because the Sun is not solid, it really has no "surface." Rather, when the density and temperature of the gases that make up the Sun drop to low enough values, the gases become transparent to light, and the heat energy then shines out into space as thermal (blackbody) electromagnetic radiation. The photosphere itself is not a sharp boundary, but a rather thin zone. The distance from the center of the Sun to the top of the photosphere is generally regarded as the Sun's radius.

METHODS OF STUDY

The interior of the Sun is difficult to study. It cannot be imaged directly. However, it is important because the Sun's magnetic field is produced deep inside the Sun, and the Sun's magnetic field and magnetic behavior are responsible for solar activity, solar storms, and considerable interaction of the Sun with the rest of the solar system.

Though the Sun's interior cannot be imaged directly, it can be studied by the way that it influences the surface of the Sun. The Sun resonates in certain vibrational modes. This can be thought of as being analogous to the ringing of a bell, only at far lower tones because the Sun is so large. These vibrations can be recorded on the surface of the Sun through observations of the rise and fall of the Sun's surface using the Doppler shift of spectral lines in the Sun's spectrum. Study of the motions of the surface of the Sun in this manner is called helioseismology. As shock waves pass through the interior of the Sun, they diffract (or bend) through the different layers of the Sun. Shock waves can also reflect from different layers in the Sun. Motion of material within the Sun can also distort the shock waves. Helioseismology can therefore yield a great deal of information about the Sun's interior structure.

Solar astrophysicists have learned much about the motion of material in the Sun's convective region beneath the photosphere. The interior of the Sun has been found to rotate at a slightly different rate than the surface layers of the Sun. This is believed to play a role in the Sun's magnetic behavior.

A further probe into the Sun's interior is through the study of neutrinos streaming from the Sun. Nuclear fusion occurring at the Sun's core produces not only gamma rays but also tiny, weakly interacting particles called neutrinos. These particles then flow outward from the Sun and can be detected on Earth. Studies of those neutrinos can yield information about the nuclear fusion processes going on in the Sun's core. Early studies of the neutrinos appeared to show far fewer neutrinos than had been predicted by nuclear theory. However, more recent studies suggest that the neutrinos oscillate, or change form, between three types, and the early experiments only detected one type of neutrino.

CONTEXT

The Sun is the nearest star to the Earth, and it is the center of the solar system. It dominates everything else in the solar system. The interior structure of the Sun is determined by its composition and mass, and that interior structure determines the observational characteristics of the Sun. Understanding the interior structure of the Sun, therefore, is important to understanding the Sun itself and ultimately its interactions with its planets, including Earth.

For many years, it was assumed that because the Sun is hot and bright, like fire, it must be shining through some sort of burning process. Burning, though, was unable to account for the Sun's energy. Astronomers theorized that perhaps the Sun was shining through release of its gravitational energy. That, too, failed to account for the Sun's energy output. Finally, by the early twentieth century, physics had advanced to the point where astrophysicists understood nuclear fusion to be the Sun's energy source.

Likewise, the interior structure of the Sun could not be understood until physicists understood the physics of materials under conditions such as exist inside the Sun. Theoretical astrophysicists now describe the Sun's structure through a set of equations called a solar model. Helioseismological studies match the current theoretical models to a very high degree of accuracy.

Raymond D. Benge, Jr.

FURTHER READING

Bhatnager, Arvind, and William Livingston. *Fundamentals of Solar Astronomy*. Hackensack, N.J.: World Scientific, 2005. A complete introduction to solar astronomy that does not require extensive mathematics to understand. The book emphasizes solar observations rather than solar structure, but there is a chapter on solar structure. An extensive glossary is included.

Caroll, Bradley W., and Dale A. Ostlie. *An Introduction to Modern Astrophysics*. 2d ed. San Francisco: Pearson Addison-Wesley, 2007. A textbook designed for undergraduate astronomy majors, providing an excellent and thorough description of stellar astronomy. A knowledge of calculus is assumed.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. An excellent and thorough introductory college astronomy textbook. An entire chapter is devoted to the Sun, and other chapters cover other topics in stellar astronomy.

Golub, Leon, and Jay M. Pasachoff. *Nearest Star: The Surprising Science of Our Sun*. Cambridge, Mass.: Harvard University Press, 2001. An easy-to-understand discussion at the popular science level. Most of the book is about observations of the Sun and solar activity, but there is some discussion of the Sun's interior structure.

Hansen, Carl J., Steven D. Kawaler, and Virginia Trimble. *Stellar Interiors: Physical Principles, Structures, Evolution*. New York: Springer, 2004. A graduate-level text on the astrophysics of stars. Technical; assumes a familiarity with physics and calculus.

McDonald, Arthur B., Joshua R. Klein, and David L. Wark. "Solving the Solar Neutrino Problem." *Scientific American* 288, no. 4 (April, 2003). An extremely good review of the research that finally led to resolving the early confusing results of solar neutrino observations. A brief history of solar neutrino studies is included.

Zirker, Jack B. *Sunquakes: Probing the Interior of the Sun*. Baltimore: Johns Hopkins University Press, 2003. A very good survey of the science of helioseismology and how it is used to study the solar interior. The book is well documented and thorough without being too technical for laypersons to follow.

_______. *Total Eclipses of the Sun*. Expanded ed. Princeton, N.J.: Princeton University Press, 1995. A very good survey of solar structure written at a level easily accessible to the layperson. This is one of the best books on solar structure for the nonspecialist. Includes numerous photographs.

See also: Coronal Holes and Coronal Mass Ejections; Electromagnetic Radiation: Thermal Emissions; Infrared Astronomy; Nuclear Synthesis in Stars; Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Magnetic Field

Category: The Sun

The solar magnetic field constitutes a powerful force throughout the Sun, causing flares, prominences, sunspots, and a host of other phenomena. Charged particles from the Sun and the solar magnetic field interact with the terrestrial magnetosphere, affecting life on Earth.

OVERVIEW

A magnetic field is produced by and associated with moving charges (electrical currents). In a simple ring current, a dipolar field resembling the field of a bar magnet results. Complex interactions of charges occur everywhere in nature, and the large-scale motion of charges is extensive and intricate. Thus, depending on the current, magnetic fields can be found almost anywhere in various strengths. (For example, Earth's dipolar magnetic field is about 0.6 gauss, rather weak compared to solar active regions with field strengths of a few thousand gauss and pulsars with field strengths of 1 trillion gauss.) Interactions between magnetic fields and charged particles result in the emission of X rays, the energizing and modulation of cosmic rays, and the acceleration of charged particles. For example, the strong solar magnetic field expels charged particles from the Sun, while the weaker terrestrial magnetic field interacts with these particles, forming a protective shield around Earth by deflecting many of them, and letting some into the upper atmosphere producing the auroras (the northern and southern lights).

On a laboratory scale, the properties of magnetic fields are well known. The solar magnetic field, however, is much larger and necessarily more complex. The solar magnetic field appears to be generated and sustained by the Sun's rotation and the turbulent convective motion of charged particles below the photosphere (the Sun's visible surface). The development of improved observing techniques and instruments, including the use of space probes and space telescopes, has yielded much information on the Sun's magnetic field and its effects.

Sunspots have been seen from Earth since at least the fourth century B.C.E., as evidenced by early descriptions of them. They appear as darker spots of various sizes on the brighter solar photosphere. The temperature of the photosphere is about 5,800 kelvins, while sunspots are cooler by about 1,000 to 1,500 kelvins, thus making sunspots about one-third as bright as the rest of the photosphere. The spots have diameters ranging from a few thousand kilometers to more than 150,000 kilometers. Generally, smaller spots last only a few days, while larger ones linger for several weeks.

The solar sunspot cycle was discovered in 1843 by Samuel Heinrich Schwabe, who observed that the average number of sunspots varies systematically with an approximate periodicity of eleven years. At the beginning of a sunspot cycle, the spots appear symmetrically at solar latitudes of about 30° to 40° north and south, gradually working their way toward the Sun's equator by the end of the eleven-year period. Then the pattern repeats during the next sunspot cycle.

In 1908, George Ellery Hale—the American astronomer responsible for establishing the observatories at Yerkes, Mount Wilson, and Palomar—observed that lines in the spectra of sunspots were split into several components. Earlier, in 1896, Pieter Zeeman, a Dutch physicist, had shown that such splitting occurs in the presence of a strong magnetic field. Hale thus established that the sunspots had intense magnetic fields (up to 5,000 gauss). Since the Sun is a sphere of hot gases, such magnetic fields can exist only as a result of powerful convective currents. Hale and his colleagues at Mount Wilson Observatory tried to measure a general solar dipolar magnetic field, similar to that of Earth (whose magnetic field resembles that of a huge bar magnet), but it was not until 1953 that Horace Babcock, using a specially designed solar magnetograph, succeeded. He showed that a weak, periodically varying dipolar field of 2 to 7 gauss exists over the entire Sun.

The solar magnetic field changes over a twenty-two-year cycle, twice the eleven-year periodicity of the sunspot cycle. Sunspots appear either in pairs or in clusters, but they are always organized in coherent pairs of opposite magnetic polarity. During a particular eleven-year sunspot cycle, for example, the leading members of the sunspot pairs may display positive polarity in the Sun's northern hemisphere and negative polarity in the Sun's southern hemisphere. At the end of the eleven-year sunspot activity cycle, the polarities in the two hemispheres reverse. The magnetic cycle then repeats after two sunspot cycles.

The Sun's diffuse, dipolar magnetic field also undergoes a reversal of polarity with a periodicity of twenty-two years. A positive magnetic pole appears around the solar north rotational pole at the peak of every other sunspot cycle, while a negative magnetic pole is found around the solar south rotational pole at that time. The polar fields subsequently expand to latitudes of 50° to 60° while the sunspot numbers decline. After the sunspot minimum passes, the increase in number of sunspots with their strong local magnetic fields forces the global field back toward the rotational poles, there to coalesce and form opposite poles from those of the previous cycle.

The sunspot cycle and the twice-as-long magnetic cycle are more or less successfully explained by the solar dynamo theory, which involves interactions between the convective motions of plasma (ionized gas) beneath the photosphere and the Sun's differential rotation. Energy produced by nuclear fusion in the Sun's central core is transported slowly outward through the deep interior by the process of radiative diffusion (repeated absorption, emission, and scattering of photons by particles). Higher up, the energy is transported by radial convection currents and eddies (transport of energy by the actual motion of matter). It is generally believed that, within these convection zones, powerful magnetic fields are generated by a dynamo process. Basically, a dynamo converts the energy of motion (kinetic energy) of an electrical conductor into electromagnetic field energy. The ionized matter (plasma) within the solar convection zone is a good conductor of electric current, with the ability to retain and "freeze" the magnetic field. The Sun rotates differentially, with a rotation period at the equator of about 25 days, increasing to 30 days above 60° latitude. This differential rotation modifies the magnetic fields generated by dynamo action in the convection zone.

According to the dynamo theory advanced by Babcock, Eugene Parker, and others, the solar magnetic cycle starts with the Sun's global, dipolar magnetic field lines of force extending from pole to pole and frozen into the plasma a few hundred kilometers beneath the photosphere. The greater rotational speed at the lower latitudes means that the frozen-in magnetic field lines are carried faster and thus farther than those at higher latitudes. In due course, after many rotations, the magnetic field lines become tightly wrapped around the Sun, producing an intense east-west azimuthal magnetic field parallel to the equator. Adjacent magnetic field lines tend to repel one another, and when the field strength is a few thousand gauss, the repulsive force is strong enough that the field lines burst through the solar surface and loop back to reenter at a neighboring point, thus creating sunspots of opposite polarity and intense magnetic fields. The temperature around sunspots is lowered (dropping about 1,000 to 1,500 kelvins) as a result of lowered pressure in the region, caused by adjacent field lines repelling each other.

The dynamo theory successfully explains the reversed polarities of leading sunspots in terms of the opposite directions of the azimuthal field lines in the two hemispheres during a given sunspot cycle. The local azimuthal field gradually neutralizes the general dipole field, eventually reversing the polarity and starting the next sunspot activity cycle. The polarities, both of the dipole field and of the leading sunspots, are reversed from those of the previous activity cycle, so this starts the second half of the magnetic cycle. After two sunspot cycles, the magnetic cycle repeats in polarity.

Sunspot pairs occur as a result of random bursting of magnetic field lines upward from within the photosphere and their reentering at a neighboring point to resume their path around the Sun. Each sunspot has a central darker umbra, with a temperature of about 4,500 kelvins, surrounded by a lighter penumbra, with a temperature of about 5,500 kelvins. Most sunspot umbrae have diameters ranging between 4,000 and 22,000 kilometers, and magnetic field strengths ranging between 2,500 and 3,500 gauss. Smaller sunspots, with umbrae ranging from 1,400 to 3,600 kilometers in diameter and field strengths averaging 2,000 gauss, are known as "pores." Compact magnetic structures with diameters less than 1,000 kilometers, field strengths less than 1,500 gauss, and no discernible penumbral region are called "magnetic knots." The lifetime of the sunspot and its associated magnetic field typically is days to weeks, and in general it is found to be proportional to the total field strength of a sunspot or an active region.

Large clusters of magnetically active regions tend to produce prominences and flares because of excessive magnetic buoyancy. Prominences are relatively cool masses of gases arching above the photosphere into the corona, following magnetic lines of force. According to the Skylab data of 1973, eruptive (or active) prominences, triggered by large-scale magnetic fields associated with major sunspot activity, widely dissipate the field and disperse charged particles out into interplanetary space. Flares (even more violent bursts of energetic charged particles and electromagnetic radiation) occur near complex groups of sunspots, and in the process, intense magnetic field lines emerge and dissipate into space.

When high-energy charged particles from large flares reach Earth, they interfere with Earth's geomagnetic field and terrestrial communications systems. Coronal mass ejections (CMEs) are giant magnetic bubbles of ionized gas that carry enormous amounts of energy into space. If they encounter Earth, they can dump enough energy into its magnetosphere to cause disruptions of communication and electrical power distribution systems. Such phenomena gradually finetune the solar dynamo mechanism for the succeeding cycle. The solar wind (the continuous flow of charged particles from the corona into the far reaches of interplanetary space) carries a tenuous magnetic field with it as well. Fluctuations in the solar wind, as well as in X-ray and ultraviolet emissions from the Sun, are closely correlated with the magnetic cycle.

Evidence for long-term variability in solar activity and corresponding fluctuations in the solar magnetic field comes from the quantitative analysis of radioactive carbon-14 deposition in tree rings. Cosmic rays from various sources normally reach Earth's upper atmosphere and produce several radioactive isotopes, including carbon 14, which results from a collision between energetic particles and nitrogen nuclei. During a prolonged period of reduced sunspot activity, the lack of turbulence in the Sun's magnetic field will allow more intense cosmic rays to reach Earth, thus increasing the rate of production of carbon 14, which is absorbed by vegetation and eventually deposited in tree rings. Historical records (often indirect) show a good correlation between climate variations and various indicators of solar activity (such as sunspot numbers, auroras, and the size of the corona seen during solar eclipses). The Spörer minimum of the mid-fifteenth century (corroborated by the recorded paucity of auroras), the Maunder minimum (a seventy-year hiatus in sunspot activity, beginning about 1645), and the Little Maunder minimum between 1800 and 1830 are examples of periods of solar inactivity and colder climates on Earth.

Searches for magnetic cycles in nearby Sun-like stars indicate that in earlier times, the solar cycle may have been more irregular, erratic, and intense than it is now. Stronger magnetic fields, causing a significant loss of electromagnetic and particle energy from the Sun, may have contributed to a loss of angular momentum, thus decreasing the rotation rate. In any case, as stars age, their magnetic cycles tend to become more regular and well established. The Sun's magnetic field strength is expected to decline as it slowly enters the red giant stage. The expanding Sun's rotation rate will slow down and the mass density of its convection zone will decrease, thus reducing the dynamo effect. Chromospheric emissions triggered by the magnetic field dissipation process are expected to subside to a low level; however, studies of chromospheric emission activities of some subgiants show a curious revival of phenomena associated with their magnetic fields.

KNOWLEDGE GAINED

Clearly, there is still much to learn about solar magnetism. What is known has been gleaned by both earthbound and space-based observations from the mid-twentieth century onward. The solar magnetograph invented in 1952 by Horace Babcock and his father, Harold Babcock, marks a milestone in the observational study of the Sun's magnetic field. Observations from space by the instruments onboard Skylab and the Solar Maximum Mission of regions of the electromagnetic spectrum that are blocked by Earth's atmosphere from reaching the ground have enhanced the understanding of solar magnetic phenomena and their astrophysical implications.

The advent of new and improved magnetometers has yielded a wealth of data concerning the Sun's magnetic field. Observations and measurements of small and large magnetic features and related solar atmospheric activities have led astrophysicists to develop an elaborate dynamo theory to explain the origin and variations of the solar magnetic field. Astrophysicists have established that the twenty-two-year magnetic cycle is responsible for most, if not all, of the phenomena collectively referred to as solar activity (such as sunspots, prominences, and flares), as well as the processes that result in the solar wind and the formation of the interplanetary medium.

The Skylab data of 1973 amply demonstrate the link between the magnetic activity cycle and stress-relieving eruptive phenomena such as prominences, flares, and coronal emissions. Solar plasma, carrying with it extensive unstable magnetic fields, interacts with terrestrial magnetic fields to affect the size and shape of Earth's magnetosphere and processes occurring in it. Data from the Solar Maximum Mission showed modulations in the output of solar energy attributable to the time-varying, churning action of the Sun's magnetic activity cycle.

Questions have been raised concerning the magnetic stability of a convective shell dynamo such as the one envisioned for the Sun. Pointing out the difficulty of initiating in the solar atmosphere the large electric currents associated with the observed magnetic field, numerous theorists have suggested that there may be a potent, primordial magnetic field frozen within the solar core. It has been hypothesized that when interstellar nebulae with tenuous magnetic fields collapse to form stars, they can retain intense primordial fields over a relaxation time of some 5 billion years, which is the present age of the Sun.

CONTEXT

Almost half a century after Hale's discovery of the strong magnetic fields associated with sunspots, Babcock and Babcock, in 1952, devised the modern solar magnetograph, opening up a new era of solar observational physics. Knowledge of the Sun's observed differential rotation, inferences about the Sun's convective zone drawn from computer models of its interior, and the laws of electrodynamics from physics were combined to develop the theory of the solar dynamo to explain observed magnetic phenomena.

While the larger pieces of the puzzle appear to be in place, many missing elements blur the picture. The solar magnetic field is not uniform, but filamentary or tubular. The precise manner in which the large-scale field produces observed emission features and expels unstable field elements, seemingly perpetuating the solar dynamo and creating the solar wind, is not well understood. It is known that a large-scale solar magnetic field can lead to a variety of mechanical wave modes. These may contribute to a heating mechanism, but this notion requires further observational confirmation. Virtually all activity in the solar atmosphere (except for the granulation in the photosphere) owes its existence to and is orchestrated by the turbulent and somewhat unpredictable solar magnetic cycle, yet many astrophysically important details are not known. Numerous attempts at improving the dynamo model, including the invocation of a strong primordial magnetic field, have proved to be only marginally successful. Long-term, detailed observations of the finer magnetic elements and processes, possibly through remote sensors, will be required, along with concurrent theoretical refinements.

An area of immediate and compelling interest is the further study of the long-term effects of solar activity cycles on terrestrial weather patterns and climate trends. Cosmic rays and their modulation by magnetic fields play an important role in genetic mutation and the evolution of life in general. Quantitative observation in this area will aid in evaluating the extent to which the deflecting action of the solar magnetic field has affected life. Further probing in this area may bring to light relationships between the evolution of life on Earth and the solar magnetic cycle.

Finally, research on the Sun's magnetic field has added to our insight into stellar magnetic fields in general. It has been firmly established that the Sun's magnetic cycles, flare phenomena, and coronal properties are common at least among main sequence stars similar to the Sun. For astronomers, then, the Sun is a cosmic laboratory, a window to the realm beyond the solar system.

V. L. Madhyastha

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. This well-written introductory textbook for college astronomy courses offers a good description of the Sun's magnetic field and its effects.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Provides descriptions of infrared emissions and how the Sun produces them.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. This introductory astronomy textbook contains a concise summary of the Sun's magnetic field and its effects.

Gibson, Edward G. *The Quiet Sun*. NASA SP-303. Washington, D.C.: Government Printing Office, 1973. Written by Skylab astronaut Gibson, this volume summarizes solar physics as understood at the time of the Skylab mission, addressing the solar activity cycle and myriad details involving the magnetic cycle and its effects. Aimed at solar researchers, but useful as a supplement for college and even high school students as well.

Jordan, Stuart, ed. *The Sun as a Star*. NASA SP-450. Washington, D.C.: Government Printing Office, 1981. This collection of review articles by noted authorities, although directed toward specialists in the field, incorporates a variety of observational results and theoretical models, showing clearly the complexities involved in understanding the ever-changing pattern of solar magnetic fields. With its numerous references at the end of each article, the volume serves as a valuable reference work on the subject at an advanced college level.

Newkirk, Gordon, Jr., and Kendrick Frazier. "The Solar Cycle." *Physics Today* 35 (April, 1982): 25-34. Using the solar magnetic-dynamo model, the authors discuss eleven-year and twenty-two-year solar cycles. The Maunder minimum in the seventeenth century and other historic periods of low sunspot activity are considered. Discussions of a possible clock in the Sun, the variation of luminosity correlated with sunspot activity, and photospheric pulsation are included, along with a list of useful references.

Parker, E. N. "Magnetic Fields in the Cosmos." *Scientific American* 249 (August, 1983): 44-54. The author, an authority on the subject of cosmic magnetic fields, presents the theory of the solar dynamo, the mechanism believed to be the source of magnetic fields of the Sun, the planets, and the galactic plane. The same mechanism, combined with the differential rotation of the Sun with latitude, affords a natural explanation of the intense magnetic fields associated with sunspots, flares, and other aspects of the sunspot cycle.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A college textbook for introductory astronomy courses. Contains several sections on the Sun's magnetic field and its effects.

Walker, Arthur B. C., Jr. "Golden Age for Solar Physics: New Instruments for Astronomy." *Physics Today* 35 (November, 1982): 60-67. An overview of solar physics and of projected lines of investigation of the multitudes of problems remaining to be solved as of the early 1980's. Features a clear account of the solar activity cycle, magnetic explosions, and phenomena triggered by the magnetic field: flares, coronal loops, and solar winds. Includes a list of references.

Wilson, Olin C., Arthur H. Vaugan, and Dimitri Mihalas. "The Activity Cycle of Stars." *Scientific American* 244 (February, 1981): 104-119. After discussing the Sun's eleven-year cycle, the authors report on the ongoing study of some ninety nearby stars and the quest for similar cycles. The goal is to understand why such cycles arise in the first place and why they sometimes suddenly vanish, as in the case of the Sun, only to reappear decades later. Astronomers may eventually be able to predict the course of the solar activity cycle from systematic observation of a large number of Sun-like stars.

See also: Coronal Holes and Coronal Mass Ejections; Earth's Magnetic Field: Origins; Earth's Magnetosphere; Earth-Sun Relations; Interplanetary Environment; Jupiter's Magnetic Field and Radiation Belts; Neptune's Magnetic Field; Planetary Magnetospheres; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Uranus's Magnetic Field.

Solar Photosphere

Category: The Sun

The solar photosphere is the visible surface of the Sun. Because it is opaque, it is not possible to see the Sun's interior layers directly. Convection currents transport energy from the interior through the photosphere, and this convection, along with magnetic activity, causes surface features such as granules and sunspots. Solar magnetic activity may affect Earth's climate.

OVERVIEW

The Sun is a ball of gas and, as such, has no solid surface. The Sun's photosphere approximates its surface, marking the region between the solar interior and the chromosphere (the region that approximates the solar "atmosphere"). Because the photosphere is relatively opaque, it blocks any view of the solar interior. As a result, most images of the solar disk show only the photosphere. The photosphere is also the coolest layer of the Sun, with a temperature of about 6,600 kelvins near its bottom and a temperature of about 4,400 kelvins near the chromosphere at the top; in the 400 kilometers between, the photosphere's temperature gradually decreases with altitude above the solar interior. Astronomers know that the photosphere is relatively cool, because it displays absorption lines that are darker than the spectrum of the hot solar interior. The temperature of the overlying chromosphere is likewise hotter, climbing to about 30,000 kelvins as the distance from the photosphere increases. Where the chromosphere ends is the corona, and in the approximately 100-kilometer-thick transition between the chromosphere and corona, the temperature rapidly increases from about 6,000 kelvins to several hundred thousand kelvins. In the corona itself, the temperature can be millions of kelvins, but the gas is very thin.

Looking at a picture of the solar disk, one can notice that the edge of the disk appears darker than the central portions. This illusion is called limb darkening. In the central portion of the solar disk one sees light from the deepest and hottest layer of the photosphere. Near the Sun's edge—its limb—only the upper, cooler layers of the photosphere are visible, because the very edge of the solar disk does not have the base of the photosphere in our line of sight behind the top layers. The cooler upper layers of the photosphere emit less energy than the hotter, deeper layers, so the limb of the solar disk appears darker. It is not really dark; it is just less bright. Limb darkening, therefore, indicates that the lower layers of the photosphere are hotter than the upper layers.

Because it is cooler than the hot, compressed gas of the solar interior, the Sun's photosphere produces an absorption line spectrum (a continuous spectrum with dark absorption lines superimposed on it at certain wavelengths). In 1814, Joseph von Fraunhofer observed the solar spectrum, and—because he was unable to identify the elements producing the various spectral absorption lines he observed—labeled the prominent lines using capital letters.

The most obvious features of the Sun's photosphere are the granules, which give the photospheric surface a mottled light and dark appearance. Granules are bright regions that vary in size but are typically about 1,000 kilometers across. They also are temporary features, typically lasting from about five to thirty minutes. Granulation results from convection currents below the photosphere.

Nuclear reactions powering the Sun take place in the core, and the energy must be transferred from the core to the surface. In the deep interior of the Sun, the energy is transferred by radiation. In the upper portions of the interior, just below the photosphere, the energy is transferred by convection currents. The solar convection currents are similar to the convection currents that heat a room, if there is a radiator in one side and no fan to blow the warm air to the other side. These solar convection currents form convection cells, regions where the hot gas from the interior flows up and then flows back down after it cools. The granules that we observe are the tops of these convection cells. They appear brighter because they are still hot from the Sun's interior energy.

Supergranules are larger versions of granules. They might typically be about 35,000 kilometers across. A typical supergranule might last one or two days, much longer than the typical lifetime of a granule. Rather than being observed directly, like granules, supergranules are typically observed from Doppler maps of the Sun's surface. The upward-moving material is moving toward Earth, and hence toward the observer, so the wavelengths of common spectral lines are blueshifted to slightly shorter (higher-energy) wavelengths and hence appear brighter.

Granules are part of what astronomers call the "quiet Sun," which comprises the solar features that are always present. Features that are present or more common only during the maximum of the solar magnetic activity cycle comprise the "active Sun." Features of the active Sun found in the photosphere are sunspots and faculae.

Sunspots are dark areas on the Sun's photosphere. Sunspots appear dark because they are about 2,000 kelvins cooler than the rest of the photosphere. Although this temperature is still quite hot, sunspots are nevertheless cool when compared to the background of the rest of the photosphere. Large sunspots can extend to tens of thousands of kilometers, larger than Earth but still very small compared to the Sun's size. Sunspots will typically cover less than about 1 percent of the Sun's photosphere.

Magnetograms of the solar surface measure the magnetic field strength at points across the Sun's surface. Magnetograms show that sunspots are regions of intense magnetic fields, with magnetic field strengths up to a few thousand times stronger than the rest of the Sun's surface. Sunspots are darker and cooler than the rest of the photosphere because these strong magnetic fields conspire to deflect the convection currents bringing heat energy from the interior to the photosphere. The reduced heat flowing from the interior causes lower temperatures.

Faculae are regions on the photosphere that are hotter and therefore brighter than the rest of the photosphere. They might be thought of as the opposite of sunspots; like sunspots, faculae are regions of strong magnetic fields, but for faculae the magnetic fields concentrate, rather than deflect, the energy from the interior. Faculae form in regions surrounding sunspots and in lower boundaries between the elevated granules. When faculae extend up into the chromosphere they are called plages.

The numbers of both sunspots and granules wax and wane in an eleven-year sunspot cycle. This eleven-year cycle in the number of spots is actually half of the Sun's twenty-two-year magnetic activity cycle. Sunspots form in groups that contain leading and following spots. The magnetic polarities (north or south magnetic poles) of the leading and following spots in a group interchange each eleven-year cycle to produce a twenty-two-year magnetic activity cycle. The last sunspot maximum occurred around 2001.

KNOWLEDGE GAINED

The Sun's total energy output, or luminosity, is largely the energy from the photosphere. Hence if the photosphere's brightness changes, the Sun's total energy output and the solar radiation reaching Earth also change. Variations in the Sun's luminosity could cause, and very likely have caused, climate changes on Earth. Changes in the Sun's luminosity are so small that, from using ground-based measurements, it is difficult (but not impossible) to measure the Sun's luminosity accurately enough to measure those variations. With the advent of space-based observatories—which eliminated the need to correct for the amount of light absorbed by Earth's atmosphere—it became possible to measure the Sun's luminosity more easily and accurately. Beginning in the late 1970's, satellite data show that when there is a maximum number of sunspots, the Sun is a very small amount more luminous than at sunspot minimum.

These data, when combined with the long-term sunspot record, suggest that changes in the Sun's energy output may have caused the Little Ice Age in the seventeenth century. The Little Ice Age was a period of more than two centuries that was colder than normal and that coincided with a prolonged period of virtually no sunspot activity, from about 1645 to about 1715 C.E. There was also a prolonged medieval grand maximum in sunspot activity from about 1100 to about 1250 C.E. This period was the warmest of the past millennium. It is likely, therefore, that changes in the amount of magnetic activity have affected Earth's climate at other times in the past. It is also possible that such changes may be a contributing factor to current global warming. The Sun has been fairly active since the mid-1700's, except for minor, brief decreases in sunspots in the early 1800's and in the late 1800's.

CONTEXT

The Sun's energy is crucial to life on Earth. Without it we would not survive. Hence we want to understand the Sun and all its components, including the photosphere. As the visible disk of the Sun that we see, the photosphere is the immediate source of this energy. Changes in the photospheric energy output could have drastic effects for life on Earth and may even affect climate change.

Visible light, ultraviolet light, infrared light, radio waves, X rays, and gamma rays are all forms of electromagnetic waves. The significance of visible light is that it is the region of the electromagnetic spectrum that is detectable by the human eye. This range is determined by the Sun's photosphere. With a temperature of 5,800 kelvins, the photosphere is brightest at the wavelength region from red to blue, peaking at yellow. This is also the wavelength of peak sensitivity of the human eye and most animal eyes. Our eyes evolved to detect the wavelength region of the electromagnetic spectrum that is most plentiful.

Paul A. Heckert

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. One chapter of this very readable introductory astronomy textbook covers the Sun.

Frazier, Kendrick. *Our Turbulent Sun*. Englewood Cliffs, N.J.: Prentice-Hall, 1980. This book gives a readable account of our knowledge of the Sun through its publication date.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. All aspects of the photosphere are well covered in this excellent introductory astronomy textbook.

Golub, Leon, and Jay M. Pasachoff. *Nearest Star: The Surprising Science of Our Sun*. Cambridge, Mass.: Harvard University Press, 2001. Provides a detailed summary of our knowledge of the Sun. The section on the photosphere goes into considerable detail on sunspot activity.

Heckert, Paul A. "Solar and Heliospheric Observatory." In *USA in Space*. 3d ed. Edited by Russell Tobias and David G. Fisher. Pasadena, Calif.: Salem Press, 2006. Describes the SOHO mission, which was used to study the Sun's outer layers, including the photosphere, and helped us understand the Sun's magnetic activity and why the temperature increases in the chromosphere and corona.

Hester, Jeff, et al. *Twenty-First Century Astronomy*. New York: W. W. Norton, 2007. Chapter 13 of this very readable astronomy textbook is about the Sun. The discussion of sunspot activity and its effect on Earth is detailed yet accessible.

Morrison, David, Sidney Wolf, and Andrew Fraknoi. *Abell's Exploration of the Universe*. 7th ed. Philadelphia: Saunders College Publishing, 1995. The Sun is covered in chapter 26 of this classic astronomy textbook.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. Cambridge, England: Cambridge University Press, 2002. An extremely well-written introductory astronomy textbook. Chapter 12 is an overview of the Sun.

Zeilik, Michael, and Stephen A. Gregory. *Introductory Astronomy and Astrophysics*. 4th ed. Fort Worth, Tex.: Saunders College Publishing, 1998. Pitched at a level for undergraduate physics or astronomy majors, this textbook goes into more mathematical depth than most introductory astronomy textbooks. Chapter 10 covers the Sun, including the photosphere.

See also: Auroras; Coronal Holes and Coronal Mass Ejections; Electromagnetic Radiation: Thermal Emissions; Infrared Astronomy; Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Radiation

Category: The Sun

The total solar radiation that falls on Earth is the primary factor in determining Earth's weather and climate. Even the smallest variation in the solar irradiance, if sustained, could alter the terrestrial environment drastically.

OVERVIEW

Solar radiation spans the entire electromagnetic spectrum, from very short-wavelength, high-energy gamma rays given off by some solar flares to extremely long-wavelength, low-energy radio waves given off by magnetic disturbances associated with sunspots and other kinds of solar activity. Between these two extremes, there are X rays, ultraviolet light, visible light, and infrared radiation, all of which provide clues to the processes occurring in and on the Sun. Astronomers study solar radiation and the solar spectrum with the use of ground-based telescopes, high-flying aircraft, balloons, and spacecraft. Study of solar radiation addresses one of the most important problems in solar physics: Does the Sun change its radiation output over time, or is it constant?

The Sun's luminosity is the total electromagnetic energy emitted by the Sun per unit time into space, or the solar radiative power. Solar luminosity is approximately 3.8×10^{26} watts, meaning it radiates 3.8×10^{26} joules of electromagnetic energy into space every second. The solar radiation per unit time per unit area impinging upon the top of Earth's atmosphere is known as the total solar irradiance, also referred to as the "solar constant." This value is $1,368 \pm 7$ watts per square meter. Whether the Sun's luminosity is really constant, however, is in question. A decrease of as little as one-half of 1 percent over a period of one century could send the entire Earth into an ice age. An increase of the same magnitude could produce first a global tropical rain forest and eventually scorched desert continents.

Most of what is known about the Sun's radiation is derived from an analysis of its electromagnetic spectrum. The Sun's visible spectrum, like that of most other stars, consists mainly of a smooth distribution of intensity of emitted light known as the continuum (or the continuous spectrum), with narrow dips in brightness, termed dark absorption lines. An oversimplified explanation is that the continuum is emitted from the Sun's photosphere, an extremely thin shell of gas giving off the visible light that can be seen with the human eye (although damage to the retina is severe if the Sun is observed directly). Also known as Fraunhofer lines after their discoverer, solar absorption lines are produced when electrons in atoms in the Sun's atmosphere absorb photons with specific energies or wavelengths of light, causing the electrons to jump to higher energy levels in the atoms.

The visible colors are only a narrow band of wavelengths within the entire electromagnetic spectrum. The wavelengths of the visible light range from just under 400 nanometers in the violet to more than 700 nanometers in the far red. Approximate wavelength boundaries for the other regions of the electromagnetic spectrum outside the visible range include the near ultraviolet, from about 120 nanometers up to just below 400 nanometers, where visible violet begins; extreme ultraviolet, wavelengths between about 10 and 120 nanometers; soft X-ray wavelengths, between about 0.1 and 10 nanometers; hard X rays, between about 0.001 and 0.1 nanometers; and gamma rays, shorter than about 0.001 nanometers. At the other end of the spectrum, near-infrared emissions range from more than 700 nanometers (where visible red ends) up to about 1,000 nanometers; far infrared refers to wavelengths between about 1,000 and 1 million nanometers (about 1 millimeter, or 1,000 microns); and radio waves are longer than 1 millimeter and can be many kilometers in length. All these regions of the spectrum provide information about different layers of the solar atmosphere as well as about aspects of solar activity associated with sunspots, prominences, and flares.

The spectral distribution of the Sun's electromagnetic radiation approximates that of a blackbody, which is a hypothetical object that is opaque to all the radiation that falls upon it. It is a highly useful concept for describing the way stars radiate their continuous spectra. Blackbodies of higher temperatures emit much more electromagnetic energy at all wavelengths than those of lower temperatures, and the peak of the radiation distribution (spectrum) is at progressively shorter wavelengths for increasing blackbody temperature. The Sun's continuum radiation distribution is similar to that of a blackbody whose temperature is about 5,800 kelvins.

Changes in the Sun's output of electromagnetic energy and hence the solar "constant" could profoundly affect the climate on Earth. There is some evidence that short-term climatic variations, measured in terms of decades to millennia, may be brought on by changes in the Sun's luminosity and the solar constant. Astronomers have wondered whether even shorter-term climatic changes, such as extensive droughts, might be caused by changes in the Sun's energy output. (On the other hand, long-term variations, such as the advance and retreat of continental glacial ice sheets over periods of tens of thousands to hundreds of thousands of years, are best explained by seasonal and latitudinal variations of the solar energy input caused by changes in the geometry of Earth's orbit and axial tilt. Even longer intervals of hundreds of millions of years between major continental glaciations are thought to result from changes in ocean and continent geometries caused by continental drift.)

Claude Pouillet, who introduced the concept of the solar constant in 1837, tried to measure the solar constant by monitoring the temperature of a blackened box filled with water. The temperature increase of the water per unit of time would reflect the energy gained via sunlight. Samuel Pierpont Langley and Charles Greeley Abbot were later pioneers in measuring and monitoring the solar constant. In 1878, Langley invented a device that he named the bolometer. The bolometer measures the energy of incoming radiation, regardless of its wavelength. (Langley discovered far-infrared light while using his bolometer to study solar radiation.) Langley used the bolometer at high altitudes to try to measure the solar constant, making mathematical corrections for Earth's atmospheric extinction. His values of the solar constant are considered fairly accurate even by today's standards.

However, modern solar astronomers question the reality of the small variations Langley and Abbot detected in the solar constant and solar luminosity. Such doubts have arisen because of the lack of precision of nineteenth century equipment and the inability to estimate accurately experimental errors and uncertainties. It can be said that, if Langley and Abbot's study suggests anything, it suggests that the solar constant is indeed constant. Any small variations simply escaped the capabilities of their measurements, even though they were convinced that they detected small changes of about 0.5 percent or less.

Certain types of variations, however, have since been demonstrated, principally variations associated with the solar cycle. John Eddy conducted an exhaustive historical study, looking back many years for evidence of past solar cycles. Much of this work concentrated on eras before the invention of the telescope and thus necessitated gathering descriptions of the appearance of the solar corona during solar eclipses, accounts of sunspot observations, and descriptions of the aurora borealis.

Regular magnetic changes occur in the solar atmosphere over a period of about 11.2 years, the period of the solar cycle. Sunspots are observed to be at a maximum each time the cycle reaches its highest magnetic strength. The years 1969 and 1980 were such peak years; during the latter, the Solar Maximum Mission (SMM) satellite was launched. The solar cycle is driven by the Sun's interior magnetism. Magnetic fields are wrapped by the rotation of the Sun. The Sun is a differential rotator; that is, its rotation rate depends on latitude; its equator rotates with a period of about twenty-five days, but at latitudes of 75° north and south, the rotation period can be thirty days or longer. Magnetic field lines between the Sun's magnetic poles are dragged by the differentially rotating gas and coiled more and more tightly around the Sun. As the magnetic field lines become twisted and snarled, they arch into the upper layers of the solar atmosphere and produce a variety of solar magnetic activity, including sunspots, prominences (arching spires of gas), and flares (caused by magnetic trapping of thermal energy, which, at its breaking point, suddenly triggers a release of immense energy).

These cycles vary in intensity and are irregular in occurrence. The intensity of a cycle is measured most directly by the number of sunspots observed during the cycle. These numbers vary noticeably from one cycle to another. Several cycles have been observed to be deficient in sunspot production, and in a few cases almost no spots are produced. These "weak" cycles might be associated with a smaller luminosity and solar constant. Some solar physicists believe that the rate of solar rotation may vary from one cycle to another and that this rate may affect the solar constant. The magnetic centers of activity on the Sun are thought to interfere with the normal outflow of radiation; they can act to obstruct or reduce the flow outward and thus slightly modify the solar constant.

Peter Foukal and Jorge Vernazza found a possible correlation between solar rotation and the solar constant variation by examining the data from the experiments by Langley and Abbot. Their statistical analysis found a change of 0.07 percent every twenty-seven days in the observations. This tiny variation, occurring suspiciously in phase with the solar rotation, is accepted by some solar physicists and doubted by others. Some hold that such a small change is difficult to verify.

The period of solar rotation seems to have suffered "glitches" in the past. There is even modern evidence for small changes in solar rotation on a very short timescale, weeks or even days long. A known, but presumably very small, connection between solar rotation and the solar constant is the fact that as solar rotation carries sunspots around the Sun, their appearance on or disappearance from the earthward side of the Sun alters the solar constant very slightly, because less light is emitted by the sunspot than from a comparable area of the photosphere.

Variable solar rotation, if real on both long and short timescales, may also be an indication of other types of changes in the Sun's deep interior that could lead to changes in the solar constant. A. Keith Pierce and James C. LoPresto, using the McMath solar telescope at Kitt Peak, Arizona, reported in 1984 that the Sun's rotation does quickly change; it speeds up and slows down in a large cap around the polar regions within periods as short as a day or two.

Many astronomers and solar physicists have not been particularly interested in measurements of the solar constant over periods of time long enough to include one or more solar cycles. They contend that such experiments are difficult and expensive and that it is unlikely that significant changes in the solar constant could be documented. However, in response to the need for good solar irradiation measurements, it was decided to include an experiment for measuring the solar constant on board the Solar Maximum Mission (SMM) satellite, launched in 1980. Although the primary mission of the satellite was to study the Sun's behavior during the peak (or "maximum") activity cycle, it was also decided that monitoring any changes in the solar constant during the maximum was of utmost importance. Richard C. Willson thus devised an instrument dubbed the active cavity radiometer irradiance monitor (ACRIM) that was carried onboard SMM. Willson's measurements with ACRIM from SMM gave a value for the solar constant of 1,368 ± 7 watts per square meter. Furthermore, it has been established that large sunspots decrease the solar irradiance by a few watts per square meter for very short periods (a week or so).

KNOWLEDGE GAINED

Despite all the studies of the solar constant, attempts to verify its constancy or variability were inconclusive until the last few decades. Richard C. Willson measured the solar constant from high-altitude balloons in 1969 and again from an Aerobee rocket in 1976. To within 0.75 percent, the solar constant during that period remained unchanged. This period included the last half of a solar cycle. In 1967, 1968, and 1978 David A. Murcray of the University of Denver measured the solar constant from high-altitude balloons. He found no change from 1967 to 1968, but he detected a 0.4 percent change between 1968 and 1978.

The SMM spacecraft, launched into Earth orbit by the space shuttle in 1980, carried a diverse, sophisticated payload to study solar activity. Among the instruments on board was the ACRIM, which was designed by Willson to detect changes in the solar constant of as little as 0.1 percent. ACRIM measured the solar input during half of each 131.072-second measurement interval and a known radiation source onboard the spacecraft during the other half of each measurement interval, during which time a shutter blocked solar radiation. The solar irradiance received by Earth is proportional to the difference between the heating rates with the shutter open and with it closed. The ACRIM results indicate that the average value of the solar constant is 1,368 ± 7 watts per square meter. It decreased slightly but steadily between 1980 and 1986, leveling off in 1987 and 1988. There are indications of a possible slow period variation connected with the solar cycle of about 2 watts per square meter. It was also determined that sunspots cause very short term variations, for periods equal to the life of the spot.

Solar astronomers working at the National Solar Observatory, Kitt Peak, Arizona, will continue to monitor certain absorption lines in the solar spectrum that are known to be sensitive to stellar luminosity. They have found that the strengths of these lines closely correspond to the variability first detected by ACRIM, both in magnitude and in the temporal variation associated with the sunspot cycle.

CONTEXT

The goal of solar spectral irradiance measurements is to measure solar absolute brightness and how much and how rapidly it changes. The visible spectrum dominates the energy emission, and both it and the infrared seem to be relatively constant. In contrast, radio, ultraviolet, and X rays show large fluctuations associated with solar magnetic activity during the solar cycle.

Prior to the space age, measurements of solar irradiance had a very incomplete time record, even over one solar cycle. Many astronomers and solar physicists were not particularly interested in measuring the solar constant over long periods of time, contending that such experiments were difficult and expensive and it was unlikely that significant changes in the solar constant could be documented. Since the solar spectrum covers an enormous wavelength range, from radio waves to X rays, taking measurements of absolute brightness was, and remains, very challenging, inasmuch as different detectors and different techniques are required in each wavelength band. Over the electromagnetic range, at least five different experimental techniques are required for imagery and dispersion of the solar spectrum.

Another difficult problem in attempts to measure the solar constant was the attempt to determine how much energy Earth's atmosphere blocked before the radiation reached detectors on the ground. Earth's atmosphere filters out radiation according to wavelength; its blocking of many parts of the electromagnetic spectrum is very effective. Ozone in the upper atmosphere completely absorbs all radiation shorter than the very longest near-ultraviolet waves, including gamma rays, X rays, and most ultraviolet rays. At the lower end of the spectrum, water vapor and carbon dioxide strongly absorb much of the infrared, and water vapor and oxygen absorb radio waves shorter than about 1 centimeter. The ionosphere, a thin layer of charged particles at an altitude of about 100 kilometers, reflects radio waves longer than about 10 meters. Even those wavelengths that are not completely blocked, still may suffer some absorption and scattering as they pass through our atmosphere.

The solution to this problem appeared when it became possible to to make observations above Earth's atmosphere, and just about the entire solar spectrum has been observed using the Orbiting Solar Observatories (OSOs), Skylab, the Solar Maximum Mission (SMM), and the Solar and Heliospheric Observatory (SOHO).

James C. LoPresto

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley. A well-written college-level textbook for introductory astronomy courses. Good description of the Sun's atmosphere and energy output.

Eddy, John A. *A New Sun: The Solar Results from Skylab*. Washington, D.C.: Government Printing Office, 1979. Eddy highlights what was learned about the Sun as a result of the Skylab mission and the use of the Apollo Telescope Mount. Spectacular photographs make clear the enormous advances in knowledge of the Sun that were made possible by Skylab. Suitable for general audiences.

Foukal, Peter. *Solar Astrophysics*. 2d rev. ed. Weinheim, Germany: Wiley-VCH, 2004. A detailed look at the Sun and our understanding of it. Also covers the history of solar astrophysics. Suitable for undergraduates.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Good description of the Sun's atmosphere and energy output.

Frazier, Kendrick. *Our Turbulent Sun*. Englewood Cliffs, N.J.: Prentice-Hall, 1980. This book summarizes the research and observational discoveries made about the Sun between about 1950 and 1981. When possible, the author relates these developments to the impact the various solar phenomena have on Earth; for example, he discusses the question of whether the solar activity cycle affects the climate of Earth.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook. Thorough and well-written. Good description of the Sun's atmosphere and energy output.

Gibson, Edward G. *The Quiet Sun*. Washington, D.C.: Government Printing Office, 1973. This excellent survey of solar physics details what is known about the Sun, starting with the solar interior and working outward through the atmosphere, giving a chapter each to discussions of the photosphere, the chromosphere, and the corona. An informative overview is provided in the first chapter. Solar activity and solar terrestrial relations are covered in the latter chapters. A somewhat technical presentation.

Giovanelli, Ronald. *Secrets of the Sun*. Cambridge, England: Cambridge University Press, 1984. An overview of modern solar physics, with the aid of spectacular black-and-white images from a number of observatories throughout the world. The presentation of the material highlights the physical nature of the various atmospheric components of the Sun, the photosphere, the chromosphere, and the corona.

Nicolson, Iain. *The Sun*. New York: Rand McNally and Co., 1982. This atlas of the Sun includes a historical overview of solar research, explores the relationship of the Sun to the stars, the Galaxy, and the universe, and provides a detailed description of the solar interior, the solar atmosphere, solar activity, and solar-terrestrial relationships. Solar flares, the solar neutrino problem, coronal holes, solar oscillations, solar wind, and even the pragmatic problems of solar energy are discussed.

Noyes, Robert W. *The Sun: Our Star*. Cambridge, Mass.: Harvard University Press, 1982. This excellent book reviews the research done in solar physics from ground-based telescopes and space satellites. Shows how scientists' knowledge of the Sun relates to their knowledge of stars in general. Studies of the solar spectrum are shown to have provided valuable information about the complex solar atmosphere. The contents are arranged according to regions or zones within the interior and atmosphere of the Sun. Much attention is given to solar activity and its relationship to Earth.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses. Divided into lots of short sections on specific topics. Has several sections on the Sun's atmosphere and energy output.

White, Oran R., ed. *The Solar Output and Its Variation*. Boulder: Colorado Associated University Press, 1977. This highly technical volume details what was known about the measurement of the entire solar spectrum, from gamma rays to radio waves, as of the time of its publication. In addition to the measurement of the solar constant, the book takes up solar variability and the sunspot cycle. Gives ample attention to the effects on Earth's atmosphere and Earth's climate. Primarily aimed at technical audiences.

See also: Auroras; Coronal Holes and Coronal Mass Ejections; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; Infrared Astronomy; Interplanetary Environment; Nuclear Synthesis in Stars; Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Radio Emissions

Category: The Sun

Specific solar phenomena lead to the production of radio waves that affect Earth. These can be broken down into several types of radio emissions, many of which appear to be more closely tied to certain solar events than others. Detecting and understanding these emissions are important to both human activities and life on Earth.

OVERVIEW

In 1942, investigations began to determine what was "jamming" British radar and allowing German battle cruisers to pass freely through the English Channel. J. S. Hey attacked this problem. Investigations led to recognition that the Sun, and not any German technology, was the source of this kind of radio interference. Hey learned that "jamming" happened during daytime and was worse when British receiving instruments were pointed in the general direction of the Sun. Hey also discovered that the Sun at that time sported a large group of sunspots. Indeed, solar radio waves were creating the interference.

It is necessary to understand the basic structure and dynamics of the Sun before seeking to understand why radio regions are produced. Two main solar processes emit energy. Beginning closest to the core, radiative diffusion is at work. Photons are created in the center of the Sun and move outward as a result of absorption and reemission by atoms and electrons that make up the interior of the Sun. After they move out past the radiative zone, they undergo the process of convection. Convection is caused by temperature differences that make hot and cool fluids circulate. Hotter gas rises, and cooler gas falls. As cooler gas sinks closer to the Sun's core, it begins to be heated, causing it to rise again. This progression repeats itself, forming convection cells.

The innermost layer of the Sun's atmosphere is called the photosphere, and is also known as the Sun's visible "surface." Most of the light emitted by the Sun escapes through this "surface." Furthermore, sunspots are found in this region. The umbra, or dark region of a sunspot, emits approximately 30 percent less light than an equally sized area without sunspots. Above the photosphere is the chromosphere. The chromosphere is characterized by an emission line spectrum, whereas the photosphere displays an absorption line spectrum. Vertical spikes of rising gas, or spicules, as well as plages, originate in the chromosphere. The outermost layer of the atmosphere is the corona. Atoms located here are highly ionized due to extremely high kinetic temperatures. This is responsible for nonthermal radio emissions.

Variations in the Sun's magnetic field often cause explosive emissions of energy, in the form of both streams of particles and electromagnetic radiation. Visually, these can be seen as solar flares erupting from the lower corona to the photosphere, specifically where sunspots are located. However, there is more complexity to this phenomenon than this simple picture might suggest. Flares create shock waves that excite the local plasma into oscillation. These oscillations produce radio waves of a frequency equal to that of the plasma oscillations that give rise to them. Occasionally, electrons near the location of a flare will get excited and interact with strong magnetic fields. This interaction will also create intense radio emissions, called radio bursts.

Detailed studies of these solar flares were done by the radio heliograph in Culgoora, Australia. It was discovered that the explosions originated from the corona. However, there were limitations to the effectiveness of the radio heliograph. It could observe at only three radio frequencies: 327 megahertz (MHz), 180 MHz, and 80 MHz. To help overcome this problem, the Clark Lake Radio Observatory was designed to use frequencies ranging from 15 MHz to 125 MHz. This new capability unveiled radio microbursts that are believed to be caused by coronal heating. Radio emission frequencies have been discovered to be a function of electron density. Typical ranges for the lower corona are approximately 100 MHz to 1 gigahertz (GHz). Near the middle of the corona, the frequencies become 10 MHz to 100 MHz. The frequencies for the lower coronal areas are larger than higher up in the Sun's atmosphere, because electron density is greater closer to the core.

The Sun also emits radio waves through synchrotron radiation, the same process by which auroras are produced. Electrons become caught spinning around magnetic field lines and have their motion restricted in two directions: one rotating around the field line and the other in the direction along the magnetic field line. This type of radiation is more intense at radio wavelengths. When plasma expands, it can overcome the Sun's inward gravitational pull. When this occurs, it is possible for electrons to escape, creating the solar wind.

Solar radio emissions come in several types. Type I is mainly narrow-band bursts that occur frequently and can last on the order of hours to days. Type II radio bursts have a wavelength of approximately 1 meter. These bursts are usually caused by solar flares. Along with solar flares, Type II bursts can be created from highly energetic particles. They are also slow "drift bursts" that begin at high frequencies and drift to lower frequencies. Type III bursts are fast drift radio bursts because their emission frequency decreases over time. They can occur from energetic electrons escaping along magnetic field lines. Typically, these bursts have wavelengths in the meter range, but have also been discovered in the decimeter range as well. Energies of the particles that create these bursts range from 1 to 100 kilo-electron volts (keV). Bursts that have broadband qualities are Type IV and Type V. Type IV are also related to solar flares. Type V most commonly last for a few minutes but can last for longer periods of time if the frequency is decreased; this type has also been observed simultaneously with Type III bursts.

KNOWLEDGE GAINED

It was once believed that Type II radio bursts correlated with energetic solar flares, in that these bursts occur only from high-energy flares. Investigation of radio bursts, however, led to the determination that Type II bursts do in fact occur with less intense flares as well as high-intensity flares.

The robotic Cassini orbiter was outfitted with a radio and plasma wave science (RPWS) instrument, designed to measure electric fields, magnetic fields, and electron density while that spacecraft orbits the planet Saturn. Even though Cassini's main mission was to explore Saturn and its large satellite Titan, in 2003, with the spacecraft still en route to the Saturn system (Saturn orbit insertion was in mid-2004), Cassini unexpectedly picked up two Type III radio bursts coming from the Sun, one on October 28 and the other on November 4. These bursts were caused by intense solar flares. The orbiter was approximately 8 astronomical units (AU) away from the Sun at the time when it detected these flares. That these solar flares were so intense was evidenced by the detection of those radio bursts at such a great distance from the Sun. Solar flares have a huge effect on the "space weather" of the entire solar system.

Ulysses is a joint National Aeronautics and Space Administration (NASA) and European Space Agency (ESA) spacecraft devoted to studies of the Sun. Launched in October, 1990, from the space shuttle *Discovery* on mission STS-41, Ulysses was sent on a trajectory rising up above the ecliptic plane in order to fly around the poles of the Sun and thereby study the Sun at all latitudes. Its mission was to understand what is happening during solar maximum, the period in the solar cycle when a copious number of sunspots are present.

Ulysses recorded coronal mass ejections and discovered an irregular solar wind that is not very periodic. This mission also made it possible to look closely at Type III radio bursts from the Sun, because Ulysses was equipped with a unified radio and plasma wave instrument (URAP). One of the functions of URAP is to reveal the direction, angular size, and polarization of radio sources located in the heliosphere, the area of space that encompasses the Sun's magnetic field and the solar wind. Investigating radio bursts is important because it can lead to mapping out the pattern of the Sun's magnetic field by the detection of energetic particles making up these long-wavelength electromagnetic waves. Scientists have questioned how beams of particles of Type III emissions can stay together long enough to be observed. Data from Ulysses addressed this by making observations of the solar plasma's electric field. It discovered that disturbed plasma can stabilize the beams and "hold" the particles and beams together. These bunches have been termed envelope solitons. Ulysses also recorded that more shock waves occur during periods when sunspot activity increases.

CONTEXT

If properly funded, the proposed Solar Orbiter will fly near the Sun, taking measurements of radio emissions and plasma waves that can later be compared to visual measurements to make correlations. One of the main tasks of Solar Orbiter will be to further investigate interrelationships between solar flares, coronal mass ejections, and radio bursts. Solar Orbiter will have a radio spectrometer (RAS) to measure solar radio emissions of frequencies in the range of 100 kilohertz (kHz) to 1 GHz. Because of the large range of frequencies that it can observe, RAS has three component spectrometers: one ranging from 100 kHz to 16 MHz, a second from 16 to 200 MHz, and a third from 200 MHz to 1 GHz. It can observe plasma variations originating at the low coronal level of the Sun's atmosphere.

NASA's Solar Dynamics Observatory (SDO) is another orbiter designed to explore the Sun. In 2008, it was under construction at the Goddard Space Flight Center. SDO's purpose is to focus on coronal mass ejections, sunspots, and solar flares. This idea of investigating so-called space weather is important because violent weather can endanger astronauts, satellites, probes, and airplanes flying near the poles of the Earth.

Jessica Lynn Bugno

FURTHER READING

Balogh, André, Louis J. Lanzerotti, and S. T. Suess, eds. *The Heliosphere Through the Solar Activity Cycle*. New York: Springer, 2007. Investigates the Sun during the sunspot cycle to analyze correlations between the cycle and the heliosphere. Discusses theory on the subject as well as actual data from spacecraft.

Golub, Leon, and Jay M. Pasachoff. *Nearest Star: The Surprising Science of Our Sun*. Cambridge, Mass.: Harvard University Press, 2001. Gives a detailed description of the solar atmosphere, discussing properties of the photosphere, chromosphere, and corona. Provides information on how the Sun affects the inner solar system's terrestrial planets and describes future plans to study the Sun.

Hilbrecht, Heinz, Klaus Reinsch, Peter Volker, and Rainer Beck. *Solar Astronomy Handbook*. New York: Willmann-Bell, 1995. Explains solar astronomy for amateurs. Goes into detail on necessary equipment, how to build certain instruments, and what aspects of the Sun should be the focus of observations. Provides pertinent theory as well.

Kitchin, Chris. *Solar Observing Techniques*. New York: Springer, 2001. Part of the series Patrick Moore's Practical Astronomy, aimed at amateur astronomers. Includes color photographs to help the observer. Focuses mostly on visible astronomy, but some radio astronomy is mentioned as well. Also describes how to observe the Sun safely.

Lang, Kenneth R. *The Cambridge Encyclopedia of the Sun*. New York: Cambridge University Press, 2001. Begins by describing the formation of the Sun and goes into detail on properties of the Sun, as well as data on solar activity. Uses considerable amounts of physics and mathematics.

_______. *Sun, Earth, and Sky*. New York: Springer, 1997. Covers a wide range of topics on the Sun, such as composition, radiation, sunspots, and other active regions. Describes the solar wind and its effect on Earth.

Severny, A. *Solar Physics*. San Francisco: University Press of the Pacific, 2004. Describes physical processes at work in the Sun. Includes information on electromagnetic radiation generated by the Sun, from small wavelengths to the long wavelengths of radio emissions.

Stone, Robert G., Kurt W. Weler, Melvyn L. Goldstein, and Jean-Louis Bougeret, eds. *Radio Astronomy at Long Wavelengths*. New York: American Geophysical Union, 2000. Provides information on solar radio astronomy, focusing on topics that relate to low energies and long wavelengths.

Verschuur, G. L. *The Invisible Universe*. 2d ed. New York: Springer, 2007. Explains what radio astronomy is and provides a history on how it started. Historical treatment includes initial serendipitous discoveries made by pioneering radio astronomers. Covers topics such as the Sun, active galactic nuclei, nebulae, and other radio sources.

Wentzel, G. Donat. *The Restless Sun*. Washington, D.C.: Smithsonian Institution Press, 1989. Describes the structure and evolution of the Sun. Also details areas of the Sun that have much activity, such as sunspots and solar flares, as well as how some of these events have been observed. For general readers.

See also: Coronal Holes and Coronal Mass Ejections; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; Interplanetary Environment; Nuclear Synthesis in Stars; Radio Astronomy; Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Seismology

Category: The Sun

Solar seismology, or helioseismology, is the study of the oscillations that take place within the Sun. These periodic vibrations originate in the Sun's convective zone. By analyzing the motion of these oscillations, scientists can image the solar interior and develop more accurate models of the Sun.

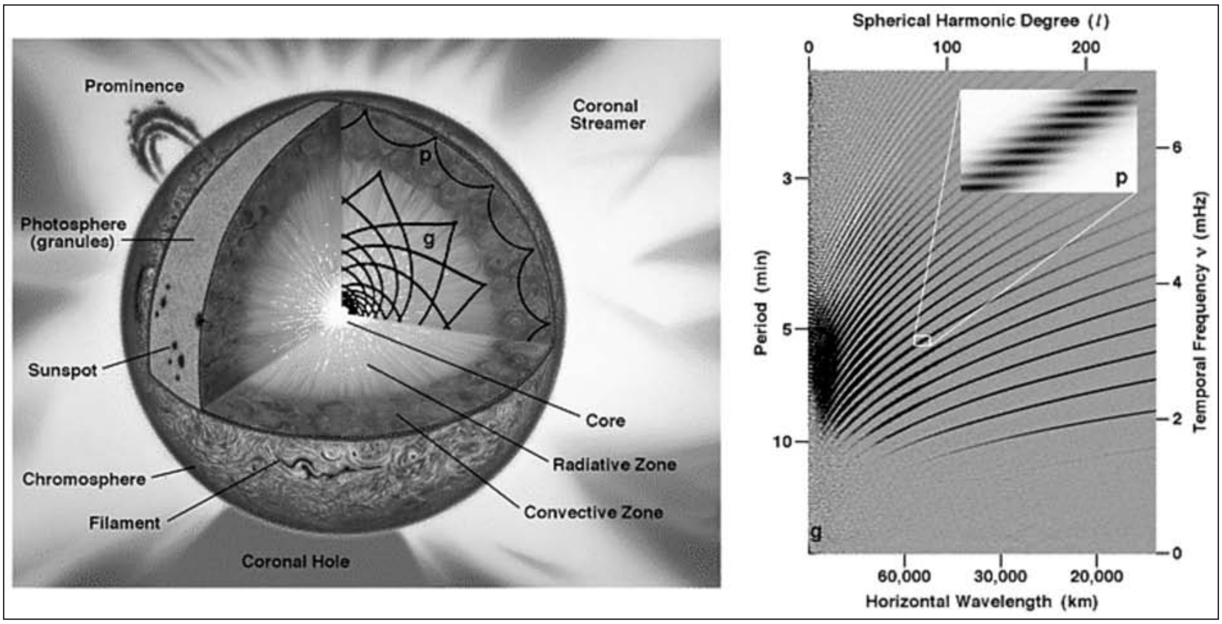
OVERVIEW

In 1960, Robert Leighton of the California Institute of Technology discovered that small areas of the Sun's surface oscillated up and down with a period of about five minutes. He observed that the spectra of various points on the Sun's surface were Doppler-shifted alternately toward shorter and longer wavelengths. The Doppler effect shifts electromagnetic radiation to shorter wavelengths (termed "blueshifts") when the source moves toward the observer and to longer wavelengths (termed "redshifts") when the source moves away from the observer. However, he was not able to tell whether the oscillations occurred only in small regions or over the entire Sun.

A few years later, it was found that the entire Sun vibrates. Like many other scientific discoveries, it happened while trying to measure something else. Astronomers were attempting to measure the oblateness of the Sun to discriminate between Einstein's general theory of relativity and a competing theory of gravity—the Brans-Dicke scalar-tensor theory. Since the mid-1800's, it had been known that Mercury's orbit around the Sun precesses; that is, the elliptical orbit itself slowly rotates about the focus occupied by the Sun. Most of the observed precession could be explained by the other planets gravitationally perturbing Mercury according to Sir Isaac Newton's laws of gravity and motion, but a small residual was unaccounted for. The existence of an unknown planet, the legendary Vulcan, inside the orbit of Mercury was suggested to provide the extra gravitational perturbation. Searches were made, and some even reported spotting Vulcan, but none of the "discoveries" was confirmed.

In the early 1900's, Einstein developed his general theory of relativity. One of its consequences is that mass distorts space-time in its vicinity. Thus Mercury's orbit would precess because of the curvature of space-time around the Sun. The amount predicted by general relativity for Mercury's precession matched exactly the observed discrepancy, and this was hailed as major evidence supporting the theory.

A half century later, Carl Brans and Robert H. Dicke, developed an alternative to general relativity called the Brans-Dicke scalar-tensor theory. It, too, predicted the precession of Mercury's orbit if the Sun were slightly oblate—that is, if the Sun had a small equatorial bulge due to its rotation. In the early 1960's, Dicke's team at Princeton University claimed to have measured a small oblateness of the Sun using a telescope of their own design and construction. A few years later, Henry Hill of the University of Arizona built a telescope that was designed specifically to detect a distortion in the shape of the Sun. He found no evidence of any distortion, but after further study, Hill and his colleagues discovered periodic oscillations of the Sun's entire surface.



These diagrams show how sound waves propagate through the interior of the Sun. Understanding this helps scientists better understand the Sun's interior structure, composition, and rotation. (NASA/ESA/SOHO)

Since then, it has been determined that the entire surface of the Sun is in a state of constant oscillation, with periods varying between minutes and hours. It might be said that the Sun is ringing like a bell being struck continuously. The origin of these vibrations is the convective zone beneath the photosphere or "visible surface" of the Sun. Bubbles of hot gas rise in convection cells, carrying heat from the Sun's interior to the surface. The tops of these convection cells produce the granulation seen in images of the photosphere. The hot bubbles rise toward the surface, accompanied by a tremendous roar. These sound waves oscillate through the Sun and cause its surface to rise and fall periodically as they are reflected by the bottom of the photosphere. As the waves travel back downward into the Sun, they encounter higher temperatures and pressures. These changing physical conditions result in the waves' velocity being increased, which eventually causes the waves to refract or bend upward toward the surface. When they reach the surface, again the waves are reflected back downward. The depth that the wave reaches depends upon its wavelength. The wavelength also determines how far a wave will travel around the Sun between reflections from the surface.

The Sun's interior is conducting waves with virtually millions of different wavelengths and frequencies. Some waves have the exact wavelength necessary to make an even number of bounces before they return to where they began. Astronomers categorize these waves by the number of times that they reflect from the surface in one complete circuit of the Sun. For example, a wave with the designation I-4 strikes the surface in three places before it bounces back to its starting position. Once it returns to its origin, it has struck the surface of the Sun four times. Scientists have found that waves with low I numbers travel deep into the Sun and can reveal physical characteristics there, while waves with higher I numbers probe the shallower zones closer to the Sun's surface.

KNOWLEDGE GAINED

On Earth, geologists use seismic waves (produced naturally by earthquakes or artificially by setting off explosions or by "thumping" the ground) as probes of the Earth's interior. Astronomers use solar seismic waves in an analogous way to image the interior of the Sun, which cannot be observed directly. Prior to this new development in solar physics, knowledge of the processes that occur within the Sun and the locations of various boundaries within the Sun came only from computer-generated models of its interior.

Currently, the generally accepted model of the Sun's interior is called the Standard Solar Model (SSM). Observed frequencies and wavelengths of observed solar oscillations for the most part are consistent with the SSM to within 0.1 percent. For example, within the Sun's core, hydrogen nuclei are fused together to form helium nuclei. This process converts a small fraction of the input mass into energy that keeps the Sun shining. The SSM indicates the central temperature needed to maintain this reaction at the right rate to account for the Sun's luminosity is about 15 million kelvins. Vibration patterns observed on the Sun's surface but that travel through the deep interior are consistent with this temperature.

Direct observation shows that the Sun's surface rotates differentially—the equatorial region has a shorter rotation period than the polar regions. Helioseismology provides a way to study the internal rotation of the Sun, and it turns out to be quite complex. At shallow depths there is a "zonal" flow with alternating bands moving faster and slower than the average. Just below the surface are wide "rivers" moving more slowly near the equator and faster near the poles, just the opposite of what the surface itself does. The base of the convection zone oscillates in rotational speed with a period of 1.3 years, sometimes moving faster than the surface and sometimes more slowly than the surface by about 10 percent. The deeper radiative zone, including the core, rotates reasonably uniformly with a period of about 27 days.

CONTEXT

The discovery that the surface of the Sun is oscillating was made in the early 1960's. At the time, scientists were gathering data on the oblateness of the Sun. As it turned out, there was no measurable oblateness of the Sun, but subsequent observations revealed an alternating Doppler shift in the solar spectra taken from various points on the Sun's surface. This Doppler shift provided evidence of periodic oscillations. Further investigations revealed that the Sun is ringing as if it were a large bell that is continuously being struck. The millions of different wavelengths and frequency combinations of waves are believed to originate within the Sun's convective zone. In this region, the tremendous heat from the interior is carried outward toward the surface by rising bubbles of hot gas. The sound waves given off by this movement of hot gases cause the Sun to vibrate.

The Global Oscillations Network Group (GONG) has provided extensive observations of solar oscillations. GONG consists of an array of 50-millimeter-aperture refracting telescopes at various locations around the Earth. These locations were selected to ensure that at least two telescopes could gather data from the Sun at all times, enabling astronomers to differentiate real oscillations, together with their amplitudes and periods, from noise produced by instrumentation or Earth's rotation. Each of the telescopes contains a Fourier tachometer. This device is capable of measuring extremely small Doppler shifts at more than sixty-five thousand different points on the surface of the Sun. By observing these shifts, astronomers can determine oscillation periods of these various points and form a detailed picture of the solar disk. This large amount of data is necessary to determine the paths that the waves follow through the Sun and convert that information into a model of the solar interior. The Solar and Heliospheric Observatory (SOHO), launched by the European Space Agency (ESA) in 1995 and on station between the Earth and Sun, also provides continuous monitoring of solar oscillations.

Helioseismic waves act as probes of the solar interior that enable solar scientists to map the interior of the Sun and solve some of the perplexing problems in solar physics. By observing oscillations visible on the Sun's surface, astronomers are assembling a detailed picture of the solar interior, in much the same way that geologists gain knowledge of Earth's interior by the study of seismic waves traveling through our planet. For example, solar oscillations provide information about the temperature and density of the Sun from its surface down to its core. They reveal the rotational speeds of internal layers of the Sun. Helioseismic imaging helps delineate the boundaries between the convective zone, radiative zone, and the core itself. Solar seismology also provides clues as to how energy is transferred from the solar surface to the chromosphere and corona. (It is currently believed that intense magnetic fields, along with acoustic shock waves from the tops of convecting cells, are responsible for high temperatures in the chromosphere and the corona—up to 2 million kelvins in the corona.) In sum, the accumulation of helioseismic data over the years is helping astronomers form an accurate model of the conditions and processes occurring both on the solar surface and in the interior.

David W. Maguire

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Includes a good description of solar oscillations, including the GONG and SOHO projects.

Cox, A. N., W. C. Livingston, and M. S. Matthews, eds. *Solar Interior and Atmosphere*. Tucson: University of Arizona Press, 1991. A series of articles covering nuclear and atomic physics topics relating to the study of the Sun. Also discusses information gathered from the world's largest solar telescopes.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Includes some discussion of solar oscillations.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook that covers what we know of solar oscillations.

Gribbin, John R. *Blinded by the Light: The Secret Life of the Sun*. New York: Harmony Books, 1991. Discusses the Sun and our understanding of it. Suitable for general audiences.

Lopresto, James Charles. "Looking Inside the Sun." *Astronomy* 17 (March, 1989): 20-28. Discusses the origin of the study of helioseismology and its possibilities in solar research. Although very little mathematics is used, the article contains an abundance of technical terms, and readers need a background in basic physics and astronomy.

Mitton, Simon. *Daytime Star: The Story of Our Sun*. New York: Charles Scribner's Sons, 1981. A nontechnical volume accessible to the general reader. Mitton discusses the Sun, its structure, its processes, and its future.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A thorough college textbook for introductory astronomy courses, including a short discussion of solar oscillations.

See also: Coronal Holes and Coronal Mass Ejections; Earth-Sun Relations; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Structure and Energy; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots.

Solar Structure and Energy

Category: The Sun

Advances in both theoretical physics and astronomical observing capabilities have provided an increasingly detailed description of the Sun. As the nearest star to us, the Sun can also function as a laboratory for investigations of stellar physics in general.

OVERVIEW

The Sun is composed mostly of hydrogen. When the Sun first formed, there was enough hydrogen in the estimated size of its core for this reaction to keep the Sun shining for approximately 10 billion years. Since the Sun is about 4.5 billion years old, it is only about halfway through with generating energy in its core by hydrogen fusion.

The Sun's mass is about 2×10^{30} kilograms. It consists primarily of the two simplest atoms: hydrogen (about 91 percent by number of atoms, about 71 percent by mass) and helium (about 9 percent by number of atoms, about 27 percent by mass). Other chemical elements account for a combined total of about 0.1 percent by number of atoms or 2 percent by mass. More than sixty chemical elements have been identified in the Sun's spectrum, and probably all the elements are present in minute amounts. Some of the more abundant of the other elements include oxygen, carbon, nitrogen, silicon, magnesium, neon, iron, and sulfur.

A major clue to the conditions inside the Sun is that it seems to be reasonably stable, with no large, rapid changes. Thus it must be in a state of hydrostatic equilibrium. This means that its self-gravity, which tries to make it contract, is balanced by its internal pressure, which tries to make it expand. If it were not in hydrostatic equilibrium (or at least very, very close to it), the Sun would either be contracting or expanding noticeably.

The Sun's luminosity—that is, the rate at which it emits electromagnetic energy—is approximately 3.8×10^{26} joules per second. The nuclear reaction that supplies the Sun's energy needs is the fusion of four hydrogen nuclei into one helium nucleus. In the production of this single helium nucleus, 4.8×10^{-29} kilograms of mass are converted into 4.3×10^{-12} joules of energy. Thus, to generate the Sun's luminosity (3.8×10^{26} joules per second), every second about 36×10^{37} hydrogen nuclei (with a total mass of 602,300,000 metric tons) are fused into 9×10^{37} helium nuclei (with a total mass of 598,100,000 metric tons), and the excess 4,200,000 metric tons of mass is converted into energy in the Sun's core.

From the center outward, the Sun has several layers or zones: the hydrogen fusion core, the radiative zone, the convective zone, the photosphere, the chromosphere, and the corona. However, other than the photosphere, these layers are either difficult or impossible to observe. Above the photosphere, the gas is so tenuous that it emits much less visible light. Below the photosphere, the gas is opaque—sort of like fog—so electromagnetic radiation cannot escape. Consequently, the layers below the photosphere must be studied indirectly via computer models of the Sun's interior. The layers above the photosphere can be seen visually on special occasions or observed anytime at various wavelengths outside the visible range.

The energy generated in the Sun's core initially makes its way outward by radiation—that is, as a flow of photons. However, because the Sun's interior is so opaque, a photon travels only about 1 centimeter before it is absorbed and reemitted by an atom. This occurs over and over again, so this method of energy transport is called radiative diffusion. The region of the Sun in which energy is transported entirely by radiative diffusion (or simply radiation) is its radiative (or radiation) zone, and this zone extends from the center out to about two-thirds of the Sun's radius.

Beyond this distance, the gas becomes even more opaque, making it more difficult for photons to travel. As a result, convection sets in, and energy is carried by rising bubbles of hot, lower-density gas. After transferring their heat, the cooler, denser bubbles sink. The region of the Sun in which convection occurs is its convective (or convection) zone, and it occupies the outer third of the Sun's radius. Even in this zone in which convection occurs, some energy is also carried by radiation, that is, by photon flow.

The next layer, the photosphere or visible surface of the Sun, is the lowest level of the solar atmosphere, with a radius of about 700,000 kilometers and a thickness of about 500 kilometers. Temperatures decrease from about 6,600 kelvins at its base to about 4,400 kelvins at its top; the overall effective temperature of the photosphere is 5,800 kelvins.

When examined with high-resolution imagers, the photosphere reveals a wealth of structure and detail. Most pronounced is the presence of granulation, an alternation of brighter spots with darker borders, resembling a mixture of salt and pepper. Each granule is a region of gas about 1,000 kilometers in diameter, larger than the state of Texas. The photosphere is at the top of the Sun's convective zone, the part of the Sun's interior where energy is transported outward by convection. The brighter region at the center of the granule is a hot gas bubble rising, while the darker border regions of a granule are cooler gases sinking back down. The convection zone in the Sun produces waves of thermal energy that shoot up through the photosphere. These waves make the photosphere appear to oscillate, with periods ranging from minutes to hours.

Above the photosphere are two more layers of the solar atmosphere, called the chromosphere and the corona, that usually cannot be seen with the unaided eye. The photosphere emits so much light and is so much brighter than the rest of the solar atmosphere that usually it overwhelms the weaker visible light from the layers above it, making them difficult to see. During a total solar eclipse, however, the Moon blocks the bright photosphere, and the fainter chromosphere and corona are visible to the unaided eye, extending out around the silhouette of the Moon.

Chromosphere literally means "color sphere." When seen during total solar eclipses, it appears as a narrow red ring just beyond the Moon's silhouette. The red color is one of the wavelengths (656.3 nanometers) that can be strongly emitted and absorbed by hydrogen atoms. The thickness of the chromosphere is about 2,000 kilometers, with temperatures of 4,500 kelvins at its base and rising to 30,000 kelvins at its top. These temperatures result in strong ultraviolet emission, and the chromosphere can be observed any time (not just during eclipses) at ultraviolet wavelengths. However, our atmosphere is opaque to most of the ultraviolet, so ultraviolet observations must be conducted by spacecraft above our atmosphere.

The chromosphere contains hundreds of thousands of thin spikes, called spicules. These spicules are jets of hot gas, hundreds of kilometers across and thousands of kilometers tall. They rise dramatically and then fall over a lifetime of several minutes, thus creating a dynamic appearance, like the dance of many small candle flames. The chromosphere also has a granulated structure. This structure cannot be directly observed but has been deduced from spectroscopic studies of the motions of gas in the chromosphere using the Doppler effect (a change in the frequency and wavelength of electromagnetic radiation caused by the motion of its source toward or away from the observer). Such studies have revealed that the chromosphere contains large, organized cells of gas called supergranules that move in unison under the influence of convective forces. These supergranules are 30,000 kilometers in diameter and contain hundreds of normal granulation regions.

Above the chromosphere, the density rapidly decreases and the temperature abruptly increases to about 1 million kelvins in a thin layer called the transition region. Beyond is the corona, the outermost part of the solar atmosphere, The corona's density is so low that it emits relatively little visible light, but during a total solar eclipse, the corona can be seen as a broad, glowing white halo of light out to distances of several solar radii. The temperature of the corona is about 1 million to 2 million kelvins. Consequently, it strongly emits X rays. It can be observed at any time at X-ray wavelengths, but our atmosphere is opaque to X rays (just as it is opaque to most ultraviolet radiation), so X-ray observations of the corona must also be conducted from spacecraft above our atmosphere.

APPLICATIONS

Our Sun is a star, the only star in our solar system. It is often described as an "average" star. While it is average in the sense that it is a "main sequence star" and just about in the middle of the ranges of stellar luminosity, mass, diameter, surface temperature, and various other properties, it is not typical in the sense that the vast majority of stars are smaller, cooler, and less luminous than the Sun. Furthermore, the comparatively small number of stars that are larger and more luminous than the Sun are so bright that they account for most of the light emitted within a galaxy of billions of stars.

Nevertheless, the Sun is the closest star to us—at a distance of about 150 million kilometers (which defines one astronomical unit, or 1 AU)—and as such provides a "laboratory" where normal stellar processes can be observed and studied at relatively close range. The next closest star, Proxima Centauri (which probably is a member of the Alpha Centauri star system and thus sometimes is called Alpha Centauri C), is about 40 trillion kilometers away, or about 270,000 times the distance to the Sun. Light from the Sun takes only five hundred seconds (a little over eight minutes) to reach Earth, while light from Proxima Centauri takes more than 4 years to reach us.

Aside from the fact the most other stars are smaller and less luminous, they do share have much in common with the Sun. Most stars have compositions similar to the Sun's—mostly hydrogen, some helium, and very small amounts of other chemical elements. Most stars, like the Sun, are reasonably stable and thus are in hydrostatic equilibrium. The light we receive from other stars, like the light we receive from the Sun, comes from a layer that resembles the photosphere (and hence other stars are said to possess "photospheres"). Observational evidence suggests that at least some stars (and probably most if not all) possess chromospheres and coronas as well—like the Sun.

Also like the Sun, stars generate energy for most of their active "lives" by nuclear fusion reactions. The fusion of hydrogen into helium in its core is the first and by far the longest-lasting fusion reaction that a star can tap. Most stars thus are generating energy the same way the Sun is, by hydrogen-to-helium fusion in their cores. After a stellar core's hydrogen is exhausted, other nuclear fusion reactions are possible, such as the fusion of three helium nuclei into a single carbon nucleus, but these higher fusion reactions do not supply energy for long.

The energy generated in a star's core is transported outward in most stars by the same two processes operating in the Sun: radiation and convection. However, computer models of stellar interiors indicate that in some stars, the convective zone is in the central part of the star, and the radiative zone is farther out. Furthermore, in some parts of some stars in some stages of their lives, a third energy-transport mechanism, conduction, is more effective than either radiation or convection.

CONTEXT

Unraveling the mystery of how the Sun and most other stars shine is one of the great achievements of science in the twentieth century. Ever since Anaxagoras in the fifth century B.C.E. asserted that the Sun was not a god but simply a big ball of fire, the question of solar fuel was a mystery. An early suggestion for the energy source was some form of chemical combustion; in other words, the Sun was literally a ball of fire that was burning something, just as Anaxagoras said. The problem was the timescale. No matter what the Sun was made of, it would have to chemically consume (burn) its entire mass in a time on the order of thousands to tens of thousands of years to keep shining at its present rate.

Another suggestion was meteoritic infall; that is, the Sun was heated by matter falling into it. Again, there is a problem of scale. A rocky asteroid 11 kilometers in diameter would have to slam into the Sun every second to release enough energy to supply its present luminosity. Also, such a rate would have produced some observable immediate effects, but none had ever been detected.

By the middle of the nineteenth century, the answer seemed to be slow gravitational contraction. The Sun would only have to shrink by a few tens of meters each year (an insignificant amount compared to its radius of 700,000 kilometers) to provide for its present luminosity. Furthermore, slow gravitational contraction could keep the Sun shining for tens of millions of years. At the time, it seemed like a reasonable duration to astronomers and physicists, but geologists, paleontologists, and biologists were beginning to figure out that the Earth and its rock and fossil record were at least hundreds of millions if not billions of years old. Again, the timescale did not fit the theory.

The accepted energy source now is nuclear fusion, a process in which light atomic nuclei are fused into heavier atomic nuclei, releasing energy. The mass of the resulting nucleus is not quite as great as the sum of the masses of the nuclei that fused to form it. That small difference in mass gets converted into energy according to Einstein's famous equation, $E = mc^2$, which says that matter, m, and energy, E, are equivalent and are related by a physical constant (the speed of light, c, squared). To be able to overcome the electrostatic repulsion that positively charged nuclei feel toward each other, nuclear fusion reactions can occur only at high temperatures (so the nuclei are moving rapidly) and high densities (so the nuclei are close together), exactly the conditions at the center of the Sun and most stars. According to computer models of the Sun's interior, the central temperature is about 15 million kelvins and the central density is about 150 times the density of water.

Our current understanding of the structure and physics of the Sun has been obtained in two ways: through observations of the Sun in all parts of the electromagnetic spectrum, and through computer models that calculate conditions inside the Sun. The processes that take place in the Sun are representative of processes that take place in stars generally. Thus our solar research has also contributed to our study of stars throughout our Milky Way galaxy and in galaxies beyond the Milky Way.

Karl Giberson

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses, with a full chapter about the Sun and its structure.

Emiliani, Cesare. *The Scientific Companion*. New York: John Wiley and Sons, 1988. A somewhat more technical book than others listed here, yet accessible to the general reader with a limited science background. Contains a full chapter dedicated to the physics of the Sun and related chapters on nuclear physics and stellar evolution.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Contains two chapters about the Sun and its structure.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. Thorough and well-written college-level introductory astronomy textbook, with a chapter about the Sun and its structure.

Noyes, Robert W. *The Sun, Our Star*. Cambridge, Mass.: Harvard University Press, 1982. A thorough treatment of all aspects of the Sun. Contains hundreds of black-and-white photographs of solar phenomena. Nontechnical and accessible to the general reader. Comprehensive and detailed.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. A thorough college textbook for introductory astronomy courses, divided into many short sections on specific topics. Three units deal with the Sun and its structure.

See also: Coronal Holes and Coronal Mass Ejections; Earth-Sun Relations; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots.

Solar System: Element Distribution

Category: The Solar System as a Whole

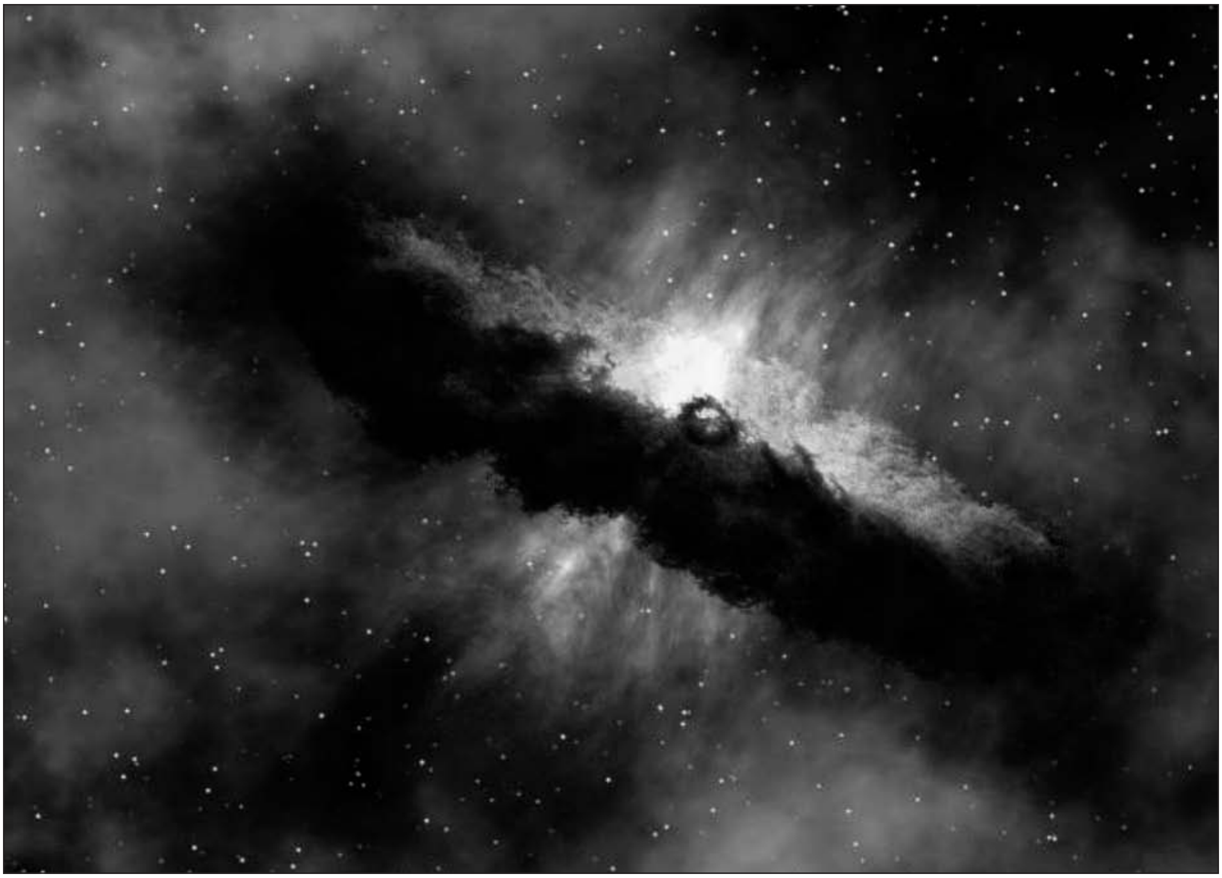
The abundance and distribution of chemical elements in the solar system resulted from many events: the creation of hydrogen and helium in the big bang, the synthesis of heavier elements by nuclear fusion reactions in earlier generations of stars, differential condensation in the early solar nebula during the accretion of the planets, and physical and chemical processes within solar-system bodies after their formation. Elemental abundance and distribution within the solar system therefore point to the origin and history of the solar system and its constituent bodies.

OVERVIEW

The solar system consists of three rather distinct parts. At the center is the Sun, composed mostly of hydrogen and helium and containing most of the solar system's mass. Next comes an inner system of four small planets (the terrestrial planets, Mercury, Venus, Earth, and Mars), three moons, and many asteroids, all composed mainly of rock and metal. Last there is an outer system of four large gas/liquid/ice planets (the Jovian or "gas giant" planets, Jupiter, Saturn, Uranus, and Neptune), many small ice-rock moons and other bodies, and lots of icy cometary nuclei. The four inner planets are made predominantly of magnesium and iron silicates, with cores rich in iron, nickel, and sulfur. The four gas giants have thick envelopes of hydrogen, helium, and hydrogen compounds in the form of gases, liquids, ices, or a combination of these, with cores of silicates and metals; these large planets generally approximate the Sun in composition. The small ice-rock bodies may have rocky silicate cores with outer layers of various ices, or they may be homogeneous mixtures of ice and rock.

A little over 91 percent of the atoms in the solar system are hydrogen, and almost 9 percent are helium; together they account for almost 99.9 percent of the atoms. In terms of number of atoms, the next two most abundant are oxygen (about 0.08 percent) and carbon (about 0.04 percent). Then come nitrogen, silicon, magnesium, neon, iron, and sulfur. Each of the other elements accounts for less than 0.001 percent of the atoms, and their abundance generally falls off rapidly with increasing atomic number. There also is a general alternation in abundance, with atoms of even-numbered elements more abundant than odd-numbered elements.

The overall abundance of chemical elements in the solar system was established by events before the solar system formed. Hydrogen and most of the helium were created in the first few minutes after the big bang; it also created a trace of lithium, even less beryllium, and possibly still less boron. Most of the atoms in our solar system heavier than helium were formed by nuclear fusion reactions in earlier generations of stars that went through their "life-cycles" before the solar system began to form. The heavier atoms were dispersed throughout interstellar space when these stars ended their energy-producing "lives," either puffing off their outer



An artist's conception of the protoplanetary disk—the cloud of interstellar gas and dust with its proto-Sun—that would form into our solar system. (NASA/JPL-Caltech)

layers as planetary nebulae or exploding violently as supernovae. The heavier atoms enriched the interstellar clouds of gas from which new generations of stars and their accompanying systems of planets formed. Measurements of the solar system abundances for the heavier nuclei roughly match the expected abundances based on the types of nuclear reactions that are expected to occur in stars.

The solar system formed when part of a large cloud of interstellar gas and dust began to contract inward under its own gravity. As the cloud contracted, it began to rotate faster to conserve angular momentum; eventually it spun off an equatorial disk of matter. Most of the mass of the contracting cloud collected in the center to form the proto-Sun, while the rest of the solar system formed in the equatorial disk. Chemical reactions governed the events that occurred in the "leftover" matter of the disk. Close to the developing Sun, only a few refractory materials condensed to form small grains of minerals and metals; far from the Sun, even gases such as ammonia and methane condensed to form grains of ice.

The first solids to condense in the early solar nebula, at temperatures around 1,600 kelvins, were oxides of calcium, aluminum, titanium, and some rare Earth elements. At about 1,300 kelvins, iron-nickel metal condensed, followed at about 1,200 kelvins by enstatite ($MgSiO_3$), the magnesium variety of the silicate mineral pyroxene. At about 1,000 kelvins, sodium, potassium, calcium, and aluminum reacted with silicate grains to create the feldspar minerals. At 680 kelvins, metallic iron reacted with hydrogen sulfide gas to create the iron sulfide troilite (FeS). Also, between 1,200 and 500 kelvins, metallic iron reacted with oxygen to create ferrous iron oxide (FeO), which reacted in turn with enstatite to create a variety of the silicate mineral olivine ($FeMgSiO_4$). Between about 600 and 400 kelvins, water reacted with calcium silicates and olivine to form various hydrous minerals. This sequence accounts for all the most common constituents of meteorites. A class of meteorites called chondrites matches the element abundance in the Sun, minus those elements which would have remained gases. Chondritic meteorites are thus considered to be the most primitive surviving materials in the solar system.

Ice sublimes quickly into gas in a vacuum. Only at and beyond the distance of Jupiter, about 5 astronomical units from the Sun, was it cold enough (about 175 kelvins) for ice to form in a vacuum and survive billions of years. At 150 kelvins, methane hydrate (a solid mixture of methane and water) condensed, followed by ammonia hydrate at 120 kelvins. Argon and pure methane solidified at about 65 kelvins.

The condensation of various solid grains at different distances from the early Sun explains much of the present distribution of materials in the solar system, and the accretion of the planets further modified their chemistry. Computer simulations of solar system accretion indicate that the small grains orbiting the Sun coalesced first into many small planetesimals (up to several hundred kilometers across) and then into protoplanets (2,000 to 3,000 kilometers across). As these bodies grew, their increasing gravity swept in matter from wider and wider swaths in the disk. The protoplanets became massive enough to attract one another, disturbing their orbits. Thus, the protoplanets swept up matter from rather wide bands in the disk, mixing materials from a variety of temperature zones. Late in the accretion process, the protoplanets collided to create larger planets. The collisions would have vaporized substantial parts of the impacting bodies, while perhaps mixing materials from two protoplanets of rather different composition.

The Earth-Moon system may be the result of such a collision. The Moon has a much lower overall density than does the Earth, suggesting that it has a lower proportion of metal to rock than does the Earth. Also, the Moon is poorer in volatile materials than is the Earth, and the more volatile a material is, the lower is its abundance on the Moon. Thus, it appears that the Moon formed in a hotter region of the solar disk than did the Earth. Many of the problems explaining the formation of the Moon can be resolved if the early Earth collided with a Mars-sized protoplanet. Much of the material of the impacting protoplanet, along with part of the Earth's crust and mantle, could have sprayed off into orbit around the Earth, later coalescing to form the Moon.

The large rocky/metallic objects possess a concentric-shell structure consisting of a dense metallic central core, a less dense silicate mantle, and a thin silicate crust. This layering reflects internal changes after the body formed. Gravitational separation of a dense metallic core occurred during or shortly after the accretion of a planet; this process may have been well under way during accretion. Collision of two large protoplanets with metallic cores might account for the unusually large core of the planet Mercury; the collision would have vaporized much of the outer rocky shells of the protoplanets, resulting in a large core but only a thin, rocky mantle and crust.

Deep within planets, high pressures crush minerals into new and denser crystal lattice structures. At depths below about 700 kilometers in the Earth, for example, magnesium silicates are crushed into densely packed cubic crystal structures. Water ice undergoes a remarkable series of changes in crystal structure as pressure increases; some of these high-pressure forms of ice are surely present in the interiors of large satellites in the outer solar system.

Internal processes may also concentrate materials on the surface of a planet or satellite to form a crust. For example, the Moon's crust formed fairly simply by melting of chondritic material. Early in the history of the Moon, its outermost 100 kilometers melted to produce a "magma ocean," probably from heat generated by the final impacts of the accretion process. As the molten material solidified, magnesium and iron atoms accumulated in dense minerals such as pyroxene and olivine, which settled to the bottom. Calcium and aluminum mostly went to form less dense feldspar, which was neutrally buoyant, neither rising nor sinking. Sinking of the olivine and pyroxene would have created a lower layer of peridotite, leaving behind an upper layer of anorthosite, matching the observed makeup of the ancient lunar crust.

When chondritic material is melted and then cooled, basalt or gabbro forms; basalt forms from rapid cooling and has small mineral crystals, while gabbro forms when the cooling is slower, allowing the mineral crystals to grow larger. Thus, it is reasonable to expect basalt and gabbro to be very common rocks in the solar system. They form the crust of the ocean basins on Earth, where they are derived from rocks of the underlying mantle. Basalt also forms the dark lava plains, or maria, on the Moon, and lava flows on Venus. The shield volcanoes on Earth are made of basalt, and similar shield volcanoes are present on Venus and Mars. The reflection spectra of some asteroids indicate that they probably have basalt on their surfaces.

In contrast to the simplicity of ocean crust, the continental crust of the Earth is chemically more complex, consisting of granite, a rock type relatively rare in the solar system; the only other possible identification of it in the solar system came from one of the Venera landers on Venus (Venera 8). Granite consists of quartz, potassium and plagioclase feldspar, micas, and amphibole. Compared with basalt and gabbro, it is greatly enriched in silica (SiO_2), potassium, and sodium, and depleted in iron and magnesium. The granitic crust of the Earth is also enriched in some less abundant elements, notably those whose atoms are unusually large (rubidium and some rare-Earth elements) and those with large atomic numbers (uranium, thorium, and lead). These elements do not fit easily into the dominant minerals of the mantle, where the principal metallic elements (iron, magnesium, and calcium) have moderate-sized atoms and atomic numbers. Granite has formed by chemical differentiation during repeated melting of the Earth's mantle and crust. Granite found on another solar-system object would be clear evidence that it had a high degree of internal activity.

Io, the innermost of Jupiter's large Galilean satellites, is a remarkable example of a body whose surface was created by internal processes. Io, though slightly smaller than the Earth's moon, is internally hot and volcanically active due to tidal flexing. Io has too little gravity to retain water vapor, the main propellant for terrestrial volcanic eruptions. In fact, Io is extremely dry, having lost most of its water to space. The only material that is cosmically abundant, volatile enough to power eruptions, and heavy enough to have been retained by such a small body is sulfur. Io has a spectacular white, yellow, and red surface, probably coated by sulfur dioxide frost and various crystalline forms of sulfur erupted onto its surface.

As protoplanets and planets grew larger, they eventually became massive enough to attract and hold gases. However, violent collisions such as those that appear to have occurred late in the accretion process would probably have driven off whatever atmospheres the protoplanets had, so smaller planets would have had great difficulty in accumulating atmospheres. In the inner solar system, the warm temperatures would have given gas molecules greater speeds, enabling them to escape more readily from small planets. Also, the early Sun probably underwent a time of intense activity (called the T-Tauri phase) during which it emitted intense streams of charged particles that swept the inner solar system free of its remaining gas. For these reasons, the inner planets have thin atmospheres or none. Earth and Venus are massive enough to have retained significant atmospheres, but Mars has only a thin atmosphere, and Mercury has only a bare trace.

In the outer solar system, the Jovian planets grew large enough to retain gases despite disruption by protoplanet collision and early solar activity. Jupiter and Saturn, which were massive enough to retain essentially all their gases, are quite close to the Sun in composition. Uranus and Neptune did not become massive enough to attract or retain hydrogen and helium quite as effectively as Jupiter or Saturn, and thus they contain somewhat less hydrogen and helium and a somewhat greater proportion of denser gases such as ammonia and methane. The heavy elements in all four Jovian planets probably accumulated into roughly Earth-sized solid cores at their centers.

Small bodies in the outer solar system accreted like those in the inner solar system, but with the addition of water ice as a major constituent. Except for Io, the small bodies of the outer solar system have silicate cores with icy mantles and crusts. Comets, whose orbits extend to great distances from the Sun, probably formed in the vicinity of Jupiter and Saturn and were expelled during close passages by the giant planets. They are mostly water ice, with other frozen gases in smaller amounts.

METHODS OF STUDY

The composition of many solar-system objects is known at least partially through spectroscopy. When materials are in the form of a diffuse gas, their atoms or molecules emit and absorb certain specific wavelengths of light and other forms of electromagnetic radiation. The specific wavelengths of electromagnetic radiation emitted or absorbed by a gas act like a chemical fingerprint of its composition. Solids have far more complex and less conclusive spectral patterns than do gases, but the spectra of light reflected from the solid surface of asteroids can be matched to laboratory measurements of the reflection spectra of various types of meteorites.

Since the generally accepted model is that the entire solar system formed during the same time from a homogeneous cloud of gas and dust, it would be expected that any solid grains that condensed from the solar nebula would have a composition similar to that of the Sun, minus those elements that would have remained gases. In the inner solar system, where the temperature of the solar nebula would have been above the freezing point of water, most hydrogen, nitrogen, and carbon would have remained gases. Most oxygen would combine with hydrogen to form water vapor, but some would combine with various metals to form oxides or would combine with silicon to form silicate minerals. Some sulfur would be available for iron-sulfide minerals such as troilite (FeS). The noble gases (helium, neon, argon, krypton, and xenon) do not chemically bond with other elements and would be nearly absent from any solid grains. This simple approximation predicts that the inner solar system bodies should consist of the most abundant elements in the Sun, minus the gases; thus the most abundant elements of the inner planets should include oxygen, magnesium, silicon, sulfur, and iron, and the actual compositions are consistent with the predictions just described.

Meteorites are particularly valuable because they provide samples of the actual materials that condensed in the early inner solar system. The type of meteorites called chondrites match the theoretical expected composition very closely and are believed to be samples of the primitive inner solar system.

Samples of rocks from the Earth's crust are easily obtained (even from ocean basins). Samples of the Earth's mantle are found in two geologic settings: as parts of ophiolite suites (fragments of displaced oceanic crust and underlying mantle) and as kimberlite pipes, or volcano-like vents that appear to have brought rocks (and occasionally diamonds) from the mantle to the surface with great speed and violence. These processes bring both shallow mantle rocks, made mostly of olivine and pyroxene, and deeper mantle rocks, or eclogites, to the surface.

Studies of the propagation of seismic waves through the Earth provide information about both the internal structure of the Earth and the physical properties of the Earth's interior. The observed physical properties match those of a magnesium and iron silicate mantle and a dense iron-nickel core.

One measure of solar system bodies can provide great insight into their chemistry, even for bodies not sampled directly. That measure is bulk density or average density, which is the mass of a body divided by its volume. The volume of a body can be readily computed once the diameter is known, and diameter can be obtained with high precision with spacecraft imagery. The same spacecraft that obtains imagery can also provide an accurate mass determination, through the gravitational effect of the body on the path of the spacecraft. Most common rocks from the Earth's crust have densities of about 2.7 to 3.0 grams per cubic centimeter. The rocks of the Earth's mantle are denser, about 3.3 grams per cubic centimeter for rocks from the upper mantle. Compression within the interior of a planet would result in higher densities, but bodies with bulk densities between 3 to 4 grams per cubic centimeter, such as the Moon, Mars, and Io, are probably made mostly of silicates of the sort found in the Earth's mantle. Denser bodies, such as Earth, Venus, and Mercury, with average densities between 5 and 6 grams per cubic centimeter, must have some denser material in addition to silicates. The most likely dense material is an iron/nickel core like that of the Earth.

The only solids that are abundant in the solar system and less dense than silicate rocks are various ices, with densities of about 1 gram per cubic centimeter. Small solid bodies with densities of about 1 to 3 grams per cubic centimeter are very likely made of varying proportions of silicate rock and ice. Most of the satellites in the outer solar system are of this composition.

The very large outer planets also have low bulk densities; Saturn has an average density of only 0.7 gram per cubic centimeter, and Jupiter, Uranus, and Neptune have average densities between 1 and 2 grams per cubic centimeter. These planets are known, from spectroscopic evidence as well as direct imaging by spacecraft, to have dense gaseous atmospheres, that, according to models of their interiors, liquefy with depth as a result of tremendous pressure. At their centers probably are solid cores of ice, rock, and/or metal.

CONTEXT

The principal value of knowing the chemical composition of the solar system is the insight it provides into its origin and evolution. For example, scientists know that the gases that formed in the early solar system did not include free oxygen. Free oxygen would have combined rapidly with other substances. Therefore, the present atmosphere of the Earth cannot be original but must have formed through chemical and biological processes on the Earth's surface.

Many elements essential to technological society, such as silver, lead, gold, and uranium, are rare in the solar system (and the universe generally). They can be extracted economically only because geologic processes on Earth have concentrated them into ore bodies. The Earth's diverse ore deposits evolved in many ways because the Earth is a dynamic planet both in its interior and on its surface. Chemical elements are concentrated in the Earth's crust by plate tectonic processes; they are further redistributed and concentrated by weathering and erosion, driven by Earth's oxygen-rich atmosphere and liquid water.

Understanding solar system chemistry also can offer a glimpse into a possible future. There has been speculation that extraterrestrial mining might someday contribute valuable resources to human civilization. Some asteroids are rich in iron and other metals; eventually it may be technologically feasible to mine these metals economically. On the other hand, the Moon lacks both the internal and surface activity of Earth, so many types of ore deposits are unlikely to form. Mining the Moon may be the most practical way to supply a future lunar colony, but it probably would involve extracting metals from common rocks, a very energy-intensive and expensive proposition. Thus it is unlikely we will mine the Moon to supply Earth.

Steven I. Dutch

FURTHER READING

Fairbridge, Rhodes W. *The Encyclopedia of Geochemistry and Environmental Sciences*. Stroudsburg, Pa.: Bowden, Hutchinson and Ross, 1972. A general reference on chemicals and chemical processes in nature, written at a moderately technical level. The article "Elements: Planetary Abundances and Distribution" deals with the chemistry of the solar system. Articles on individual chemical elements are also of interest.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college textbook beyond the introductory level, its approach is based on comparative planetology. Provides an extensive description of all aspects of the origin of the solar system and its individual members, including their composition and evolution.

Kallenbach, Reinald, et al., eds. *Solar System History from Isotopic Signatures of Volatile Elements.* New York: Springer, 2003. A collection of articles written by specialists in a variety of scientific fields, presented at the 2002 workshop held by the International Space Science Institute. They analyzed the abundance of elements in the Sun, planets, comets, and dust particles, which can reveal much about the early stages and evolution of the solar system.

Manuel, Oliver. *Origin of Elements in the Solar System: Implications of Post-1967 Observations.* New York: Springer, 2001. A compilation of works from a symposium held by the American Chemical Society. Articles were written by astronomers, geologists, chemists, physicists, and other scientists. An excellent reference work focusing solely on the chemical composition of the solar system.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. This astronomy textbook contains units that provide a thorough discussion on the structure, composition, and origin of the solar system, as well as detailed descriptions of the planets and other objects in the solar system.

See also: Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Origin; Earth-Moon Relations; Earth-Sun Relations; Eclipses; Interplanetary Environment; Kuiper Belt; Oort Cloud; Planetary Interiors; Planetary Ring Systems; Solar System: Origins.

Solar System: Origins

Category: The Solar System as a Whole

The solar system formed about 4.6 billion years ago from a cloud of gas and dust that contracted due to its own gravity. Most of the matter went to form the Sun at the center. The planets, their satellites, and the asteroids and comets formed through condensation and accretion in an equatorial disk that developed around the contracting proto-Sun.

OVERVIEW

Although many models and theories of the solar system have been generated from ancient times, most scientists today accept a version of the nebular hypothesis, which states that the solar system formed from a cloud of gas and dust. This nebula probably was part of a much larger interstellar cloud of gas and dust that contained enough matter to form several hundred to several thousand stars like our Sun. Density irregularities probably caused the larger cloud to fragment into many smaller clouds that produced stars and accompanying planetary systems. One of these smaller clouds was the solar nebula, the protosolar system.

Although the solar nebula was only a small fragment of the much larger parent cloud, it probably was at least one light-year in diameter and contained at least as much matter (and maybe up to twice as much) as the solar system today. Its composition probably was similar to what we determine spectroscopically today for the Sun's atmosphere: about 71 percent hydrogen, 27 percent helium, and 2 percent all the other known chemical elements. The hydrogen and helium were formed in the aftermath of the big bang, which occurred about 13 to 14 billion years ago. The heavier chemical elements (those up to iron in atomic mass) were synthesized by nuclear fusion reactions in the cores of massive stars, and those heavier than iron were formed when massive stars exploded as supernovae. The explosive deaths of these massive stars dispersed the heavier elements through interstellar space, where they enriched the clouds of gas and dust from which new generations of stars and their systems of planets would form.

The trigger for the gravitational contraction of the solar nebula could have been its passage through the density wave associated with one of our Milky Way galaxy's spiral arms. As the cloud orbited the center of our galaxy, passing through a spiral arm density wave would have slowed the cloud a bit, compressing the gas and starting its gravitational contraction. Support for this idea comes from observing that the spiral arms of galaxies are highlighted by groups of very young stars that have just recently formed and glowing clouds of gas that have been excited by nearby hot young stars. Another possible trigger to compress the gas and start gravitational contraction could have been a shock wave from the explosion of a nearby supernova. In fact, there is evidence based on the abundance of certain isotopes that such a nearby supernova exploded not long before the solar system formed.

At first, the cloud of gas and dust was cold and rotated slowly. Gravity pulled the material toward the center, and as the cloud contracted, it rotated faster to conserve angular momentum. It continued to contract and consequently rotate even faster, and after about 100,000 years it spun off a disk of material in its equatorial plane. Much of the matter of the solar nebula collected at the center of the disk, becoming the proto-Sun. As the proto-Sun contracted, gravitational potential energy was converted into thermal energy, heating the proto-Sun to the point where it began to shine at infrared wavelengths.

The proto-Sun continued to contract, increasing in temperature and luminosity. This heated the inner part of the equatorial disk close to the proto-Sun, but not the part farther out. The radial difference in temperature across the disk led to differential condensation into small, solid grains. At the high temperatures of the inner part of the disk, only metals and silicate minerals could condense, but farther out, where it was cooler, more abundant materials such as water, ammonia, and methane could condense into ices. The solid grains that formed by condensation collided with each other as they orbited the proto-Sun; if the collisions were not too violent, they stuck together in a process called accretion, gradually building objects up to several kilometers in size called planetesimals. Exactly how the grains stuck together is not certain; perhaps they acquired electrical charges and were held together by static electricity, or if the grains were near their melting points, they might have been somewhat "sticky."

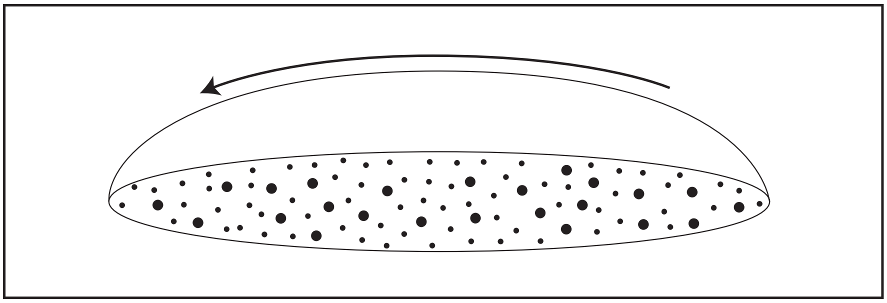
As the planetesimals grew in size, their increased mass gravitationally attracted additional material until they had swept clear the area around their orbits. The planetesimals continued to grow through collisions, becoming protoplanets. The planetesimals and protoplanets that grew in the inner part of the disk were mostly rocky and metallic in composition, while those that grew farther out in the disk were composed mostly of ices such as water, ammonia, and methane. Because water, ammonia, and methane were so much more abundant than metals and silicate minerals, the outer protoplanets grew much larger than the inner ones; with their increased mass, they were able to capture hydrogen and helium gases, which were even more abundant in the disk, and hence grew larger still. The hydrogen and helium formed thick envelopes around the planet cores, with the material in the outer part of the envelope remaining gaseous, but liquefying with depth due to increasing pressure. Eventually, only a few planet-sized bodies remained: four small, rocky and metallic inner planets (the terrestrial planets) and four large gas/liquid/ice outer planets (the so-called gas giants).

The asteroid belt, a region of small rocky and metallic bodies between the orbits of Mars and Jupiter, probably is leftover material from an early stage of solar system development that failed to coalesce into a planet, perhaps because of the gravitational influence of Jupiter. The Kuiper Belt, a region of icy and rocky bodies beyond the orbit of Neptune, like the asteroid belt probably also is leftover material that never coalesced into a major planet; Pluto is one of the largest members of the Kuiper Belt. Ices in the vicinity of the Jovian planets accreted into small cometary nuclei that, through gravitational interactions with giant planets, were tossed randomly into the far outer reaches of the solar system to become the Oort comet cloud.

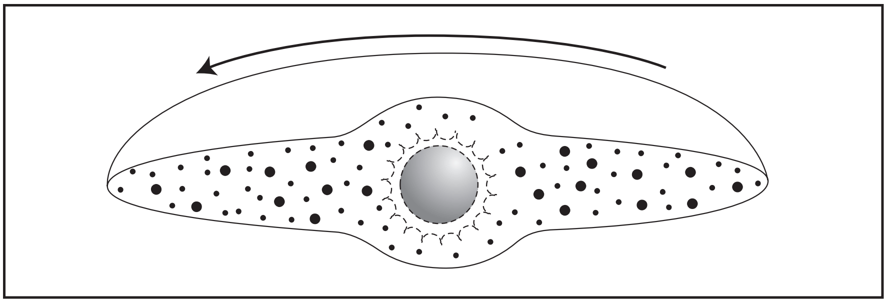
Most of the larger moons of the planets probably formed in a process similar to that of the planets themselves: by accretion in an equatorial disk around the protoplanet. Smaller moons may have formed separately from the protoplanet and were later captured gravitationally. The Earth's moon probably formed as the result of the collision of a large protoplanet with the early Earth, the impact blasting material from the Earth's crust and mantle as well as the impacting object itself into orbit around the Earth, there to accrete into our Moon.

As the proto-Sun continued to contract gravitationally, the density and temperature in its core eventually became high enough to initiate the fusion of hydrogen into helium. In this thermonuclear reaction, four hydrogen nuclei fuse to make one helium nucleus. The combined mass of the four hydrogen nuclei that go into the

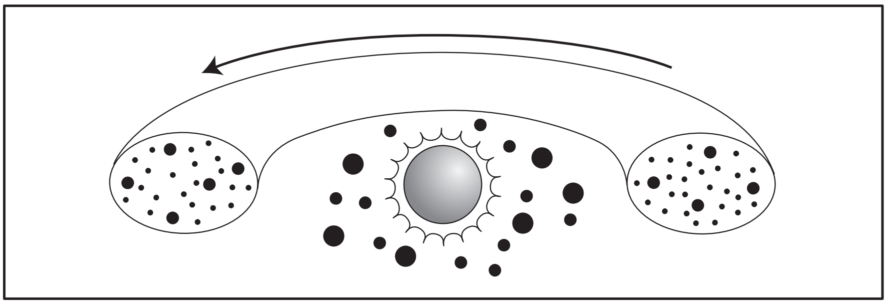
Formation of the Solar System



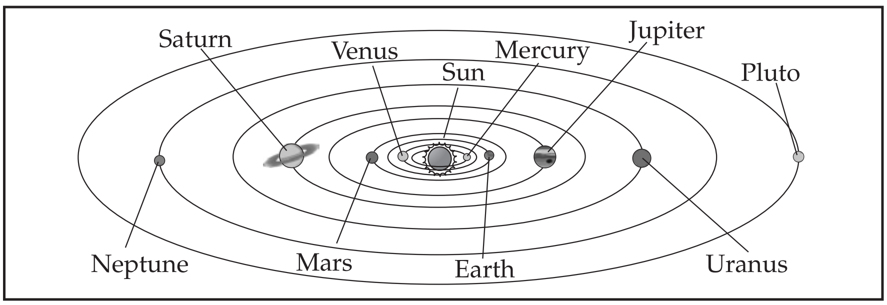
1. The solar system began as a cloud of rotating interstellar gas and dust.



2. As the force of gravity contracted the nebula, rotational speed increased to conserve angular momentum and most of the mass formed a central proto-Sun.



3. Rotation accelerated, and centrifugal force pushed icy, rocky material away from the proto-Sun. Small planetesimals formed around the Sun in interior orbits.



4. The interior, rocky material formed Mercury, Venus, Earth, and Mars. The outer gaseous material formed Jupiter, Saturn, Uranus, and Neptune. Pluto and other Kuiper Belt objects orbit beyond Neptune.

reaction slightly exceeds the mass of the single helium nucleus that results, and this small excess in mass gets converted into energy. Once hydrogen fusion was initiated in its core, the Sun became a full-fledged main sequence star.

The young Sun was much more active than it is today, shedding gases profusely from its surface into space. This became the solar wind, streams of electrically charged particles ejected from the atmosphere of the Sun. With a speed of at least several hundred kilometers per second, the early energetic solar wind blew any gas and dust remaining in the equatorial disk out into space.

The solar system took a few hundred million years to stabilize. Accretion ended with a period of heavy bombardment, when the remaining planetesimals and protoplanets, as well as icy cometary nuclei, smashed into the planets and moons; the scars left by thousands of impacts remain today on the surfaces of many of the planets and moons that are geologically inactive. It may be that much of the water on Earth came from icy cometary impacts early in Earth's history.

METHODS OF STUDY

To deduce how the Sun and planets formed, scientists are constrained by what the solar system is like today. In a sense, investigating the origins of the solar system is like a detective mystery in which ambiguous—and sometimes misleading—clues must be put together to assemble a plausible sequence of past events. Unfortunately, some clues are still missing, and still others are poorly understood.

Meteorites, asteroids, and cometary nuclei are among the smallest bodies of the solar system. These objects should have changed the least since their formation, and their composition should reflect the original material from which the solar system formed. The age of the solar system is determined by dating meteorites using the decay of radioactive isotopes they contain. Most meteorites have very nearly the same age, 4.6 billion years. The most abundant type of meteorite, the chondritic stony meteorites, also appear to be the most primitive and unprocessed. Their internal structure of small mineral grains along with somewhat larger glassy chondrules is taken to be evidence of the early period of condensation and accretion. Asteroids seem to be the parent bodies of many meteorites, based on comparison of the spectra of the light both reflect. Asteroids are thought to be examples of the planetesimals that formed through accretion, and the compositions of different types of asteroids and meteorites probably represent the processing of solar nebular material that occurred as the planetesimals grew in size. Cometary nuclei probably are unprocessed samples of the ices that condensed in the outer part of the solar nebula.

Spectra of the light emitted by the Sun and reflected by other solar-system objects provides information about the composition of their atmospheres (if any) and their surface materials. Actual samples of surface material include Earth rocks and minerals, Moon rocks brought back by the Apollo lunar missions, and a few meteorites that are thought to have been blasted off the Moon and Mars by large impacts. In addition, landers on the Moon, Mars, and Venus have sent back information on surface composition. Clues about the overall composition of solar-system objects are provided by their average density, found by dividing their mass by their volume.

CONTEXT

Just about every culture has its creation myths—stories about the origin of the world. Naturalistic explanations for the origin of stars generally and the solar system in particular can be traced back to the late 1600's, after scientists came to accept the heliocentric model of the solar system and Sir Isaac Newton published his theory of gravitation. Newton himself suggested that the Sun and stars could have formed by gravitational contraction of initially diffuse matter evenly dispersed through an infinite space. About the same time, the French philosopher René Descartes introduced perhaps the first description of what has come to be called the nebular theory. He proposed that a large cloud of gas (a nebula) had contracted under its own gravity, with the Sun forming at the center and the planets forming in the cooler outer parts. This idea was developed by the German philosopher Immanuel Kant in the mid-1700's.

The French mathematician Pierre-Simon Laplace in 1796 added conservation of angular momentum to the model, concluding that, as the nebula contracted, it would spin faster.

Alternative models for the origin of the solar system were also proposed. Many involved a second star in addition to the Sun. One hypothesis was that another star passed near our planetless Sun and gravitationally pulled matter out of it to become the planets. Another hypothesis was that our Sun was a member of a binary star system; the companion star exploded, and its debris formed the planets. However, by the early to mid-twentieth century, these alternative models fell into disfavor because of difficulties in getting them to work the way they were supposed to when physics was applied to them to try to make them more rigorous. Meanwhile, the nebular model continued to be refined to the point where today it is the generally accepted explanation.

Any successful model for the formation of our solar system must explain several patterns and regularities observed today. First, the planets all orbit the Sun in nearly the Sun's equatorial plane and in the same direction that the Sun rotates. The orbits of the planets are nearly circular. The four inner planets (the terrestrial planets) all are small, dense, and composed mainly of rocky and metallic material. The four outer planets (the Jovian planets) are large, low in density, and composed mainly of gases, liquids, and ices. The Sun, the planets, and other solar-system objects all have about the same age—about 4.6 billion years. Current versions of the nebular model account for all these points.

Solar system oddities and irregularities also need to be explained. In the basic nebular model, one would expect that the planets would all rotate in the same direction as they revolve around the Sun, and their rotational axes should be perpendicular to their orbital planes. Indeed this is the case for a few planets, like Jupiter. However, most have their rotation axes tilted at moderate angles of 20° to 30° away from the perpendicular to the orbital planes; for example, Earth's rotational axis is tilted about 23.5°. The most glaring exceptions are Uranus, which is tilted over so much that its rotational axis lies almost in its orbital plane, and Venus, which rotates very slowly and opposite to its direction of revolution. These departures are explained by impacts during the final accretion of the planets. Off-center impacts by moderately sized protoplanets could account for the moderate axial tilts of planets like Earth, and off-center impacts by large protoplanets are invoked to explain the unusual rotations of Uranus and Venus.

A feature not fully understood is the spacing of the planetary orbits; the orbits of the four inner terrestrial planets are much more closely spaced than the orbits of the four outer Jovian planets. Of course the Jovian planets are larger, which means that they formed from more material spread over a greater range of distances. Did they just happen to form this way by chance, or was it because of gravitational interactions among the planets leading to long-term orbital stability?

A major problem with early nebular models concerned the distribution of angular momentum in the solar system. The Sun, with more than 99 percent of the mass of the solar system, accounts for only 2 percent of the total angular momentum; the planets, with less than 1 percent of the mass, together possess 98 percent of the angular momentum. There had to be some mechanism for the early Sun to transfer most of its angular momentum to the equatorial disk in which the planets formed. It now generally is assumed that magnetic braking between the magnetic field of the early Sun and the ionized gas in the inner part of the surrounding disk transferred the angular momentum. Also, much of the original angular momentum could have been carried out of the system entirely by the energetic early solar wind.

Until 1995, most scientists agreed that there was only one system of planets orbiting a star: our own. Then, in 1995, the first planets orbiting other stars were detected; now several hundred such systems are known. Many of these systems are quite different from our own in terms of their planets' masses and the distances of those planets from their parent stars. Comparing the properties of our own solar system with these others can help refine our models for the formation of planetary systems generally.

Divonna Ogier and Richard R. Erickson

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. One chapter of this excellent introductory astronomy textbook provides a thorough description of the origin of the solar system.

Dermott, S. F., ed. *The Origin of the Solar System*. New York: John Wiley & Sons, 1978. A collection of papers aimed at scientists, using technical language and mathematics. Gives a broad overview of the many theories of how the solar system formed and the problems with each theory.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Part of one chapter contains a good description of the origin of the solar system.

Frazier, Kendrick. *Solar Systems*. Alexandria, Va.: Time-Life Books, 1985. A richly illustrated volume with an extremely well-written narrative designed for general readers. Concepts are explained in simple, conceptual ways and are supported by beautiful artwork, which graphically portrays the concepts.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook in which part of one chapter contains a good description of the origin of the solar system. Describes the solar system as a whole, its origin, and its future.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. A college textbook beyond the introductory level, its approach is based on comparative planetology. Provides an extensive description of all major aspects of the origin of the solar system.

Henbest, Nigel. *Mysteries of the Universe*. New York: Van Nostrand Reinhold, 1981. A well-written and beautifully illustrated volume that discusses the edges of our contemporary understanding of the solar system and the universe. Although meant for the general reader, the text assumes a basic knowledge of astronomy. While dealing with very complex subjects, it keeps within the range of the interested nonscientist.

Sagan, Carl. *Cosmos*. New York: Random House, 1980. An easily understood and nicely illustrated volume that not only describes and explains many features of the universe and the place of humankind in it but also looks at humans' relationship to the universe culturally and historically. Places the study of the universe in perspective with modern times, stressing its importance to the future.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses, divided into short sections on specific topics. Several units provide a thorough discussion on the structure and origin of the solar system.

See also: Earth's Age; Earth's Atmosphere; Earth's Composition; Earth's Origin; Earth-Moon Relations; Earth-Sun Relations; Eclipses; Interplanetary Environment; Kuiper Belt; Oort Cloud; Planetary Interiors; Planetary Ring Systems; Solar System: Element Distribution.

Solar Ultraviolet Emissions

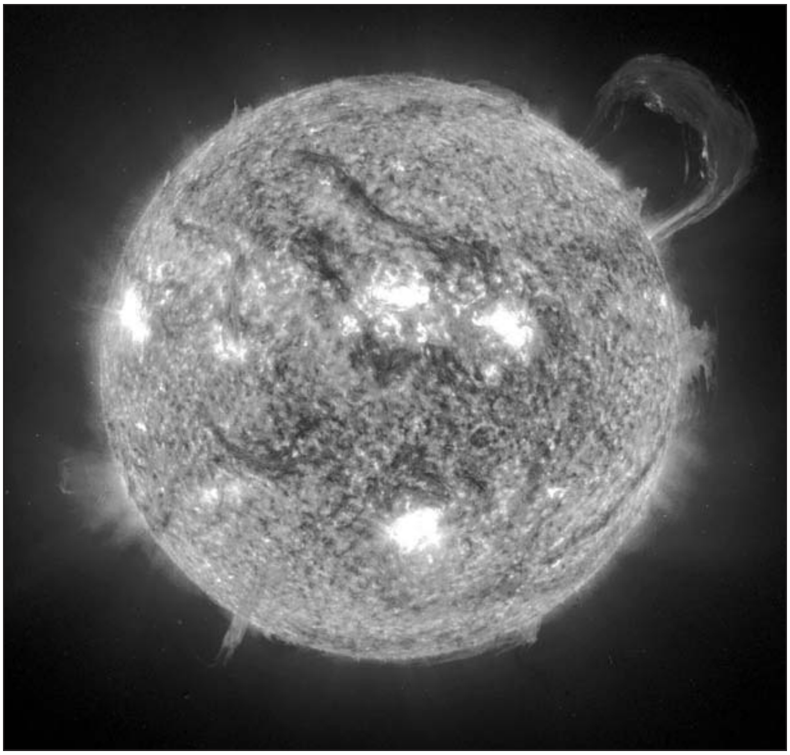
Category: The Sun

The Sun's upper chromosphere, visible as a reddish ring around the Moon during a solar eclipse, has been found to be the principal source of solar ultraviolet radiation.

OVERVIEW

Ultraviolet (or UV) radiation is a form of electromagnetic radiation between visible light and X rays. Most UV radiation is absorbed by ozone in Earth's upper atmosphere; only the "near ultraviolet" (UV with wavelengths not too much shorter than visible violet light) can penetrate to the ground. (Solar near UV is what produces suntans and sunburns.) Since Earth's atmosphere acts as an efficient UV filter, astronomical observations in most of the UV range are best conducted by instruments on board a spacecraft, beyond the atmosphere of Earth.

Observations of both UV and extreme ultraviolet radiation (EUV radiation, with short wavelengths approaching "soft X rays") are important because of the effects these emissions have on Earth's upper atmosphere and also because they can be used to model empirically the layers of the solar atmosphere known as the chromosphere and lower corona. The UV spectrum is a good spectral diagnostic for inferring physical conditions in these layers, with the brightness and variations of the spectrum providing clues about the solar atmosphere.



In September, 1999, the Extreme Ultraviolet Imaging Telescope captured this image of a large solar prominence extending like a handle from the Sun's atmosphere; the hottest areas appear as bright splotches on the surface. (SOHO/Extreme Ultraviolet Imaging Telescope Consortium)

Before the Orbiting Solar Observatories and Skylab were launched, observations of the Sun's ultraviolet region had been made by brief sounding-rocket observations, lasting for several minutes at best. The results of these observations made it clear that the study of the Sun's atmosphere and its UV radiation would be crucial to an understanding of the Sun. Observations of total solar eclipses had made solar astronomers somewhat aware that the chromosphere and corona, invisible to the human eye outside those eclipses, emit much of their light in UV and shorter wavelengths. It became apparent with the eight Orbiting Solar Observatory satellites that wavelength regions other than the visible had to be observed and mapped in detail if these layers were to be understood. It was also clear that solar EUV radiation over time underwent significant variations in brightness, variations that perhaps could be attributed to changing solar magnetic activity.

The solar atmosphere has a thin but significant layer known as the transition region, located between the upper chromosphere and corona. For a variety of reasons, the temperature, which steadily falls as the distance from the Sun's center increases outward, reverses in the upper chromosphere and spikes upward in the transition region to very high temperatures in the corona, near 2 million kelvins. Solar astronomers had long wanted to use ultraviolet images to study the upper chromosphere and transition region. The complete study of the chemical composition, temperatures, densities, and physical processes within these regions would help astronomers to understand the complex mechanisms involved in the structure and dynamics of the solar atmosphere. How is heat transferred to the corona through the transition region? How are the transition region and chromosphere related to the corona above and photosphere below?

Skylab provided the first detailed pictures of the transition region, using the ultraviolet spectrograph associated with the Apollo Telescope Mount. At UV wavelengths, the edge, or limb, of the Sun appears brighter than the disk because of the increase of temperature with increasing height above the surface. This thin region has a temperature of about 150,000 kelvins in its lower layers and 300,000 kelvins at the halfway point between the surface and the coronal layers. (This behavior is opposite to that observed in ordinary visible light, where limb darkening occurs in the visible image because of a decrease in temperature with height in the photosphere.) The duration in space and the large size of the Apollo Telescope Mount observatory telescopes enabled the astronauts to obtain sensational images that would never be possible on the ground because of interference from Earth's atmosphere. Previously, these images had not been obtainable from space either, since the weight of instrumental packages on spacecraft was highly restricted.

Observations of the UV chromosphere give insight into the way the Sun's atmosphere changes with distance from the surface, as well as details of how changes come about as a result of solar magnetic activity. The chromosphere is covered by a mesh of bright spike-like jets of gas, shooting from the chromosphere into the low corona, called spicules. When viewed near the limb of the Sun, they resemble slanted spikes arranged in uniformly spaced circles, separated by distances slightly larger than the diameters of the circles, so that they do not intermesh. This arrangement is referred to as the "spicule forest." Many solar astronomers believe that spicules play a crucial role in transferring nonradiative energy (energy produced by mechanisms other than electromagnetic radiation) into the layers of the solar atmosphere at and above the transition region.

UV images show fluffy loops that join and link active regions together in a lacelike pattern. This pattern shows up in the images as a mottled network that is no doubt directly related to the spicule network. In higher layers of the chromosphere near the transition region, this mottled network takes on a more uniform distribution.

The Skylab full-disk pictures provided a better understanding of spicules and the chromospheric network and also revealed larger spicule-like features found at the Sun's poles. These observations provided new insights into the way energy is transmitted upward to heat the outer layers of the solar atmosphere. The network fades and becomes nearly indistinguishable at the poles, because the larger spicules protrude to roughen the appearance of the solar limb. The overall chromospheric network is also distorted and made into a nonuniform pattern near centers of activity in middle solar latitudes.

Fine brushes of lines sprout from bright UV regions of the chromosphere, revealing lower legs of magnetic loops that stand in magnetically active regions. The burning-bush appearance of these curving lines is formed by the roots of the magnetic loops that correspond to concentrated magnetic areas in the photosphere where sunspots and other aspects of solar activity are observed. Above the photosphere, in the chromosphere, corresponding to higher arch structures in these same loops, other forms of activity are observed, such as prominences (large flamelike tongues of gas that appear to be extensions of the chromosphere into lower regions of the corona) and flares (sudden, triggerlike releases of tremendous amounts of thermal energy trapped by these magnetic arched loops). They appear to occur in tubelike sections of these arched loops from the corona through the chromosphere and on occasion into the photosphere. The flares are believed to occur where two or more of the arches magnetically join one another and violently rearrange their magnetic fields. This entire chromospheric ultraviolet network fades in the lower corona and gives way to a looser, less uniform distribution of material.

Almost all the light emitted by hot coronal gas comes from loops that trace out and mark patterns of magnetic lines of force rooted in bright, active regions in the chromosphere. The entire corona appears to be made up of intertwined arrangements of many of these arched loops. Snarled and twisted magnetic fields show up as particular patterns in the UV chromospheric network where magnetic activity is located; these UV patterns also signal the beginning of energetic events associated with this activity.

In addition to revealing the overall network-like structure of the chromospheric layers and the location of solar activity, the UV images from Skylab provided great insight into the structure and physical properties of various forms of transient solar activity such as prominences, surges (sudden eruptions of gas from deeper layers), and flares. Skylab obtained many detailed pictures of prominences, showing their temperature and three-dimensional structure. The ghostly prominences are shaped and supported by magnetic forces and are made up of chromospheric material elevated and immersed in the hotter coronal gas. Detailed UV images of prominences show coronal tunnel-shaped refrigerators maintained by magnetic tubes, whose outside diameter is surrounded by the hot corona. The tube has a temperature that is lowest in a line running down its center. The temperatures of these structures are inferred from the brightness of the UV image. The brightness-temperature correlation is scientifically defined in terms of a quantity called "emission measure," but simply put, the basic principle is the hotter the gas, the brighter the glow.

The ultraviolet spectrograph on board the Solar Maximum Mission (SMM) satellite gave solar astronomers the ability to observe prominences at wavelengths, and hence temperatures, not seen from Earth. These observations established the existence of downflows at the footpoints (regions of strongest photospheric magnetic fields) of the arches of the prominences. After sufficient loss of mass to lower layers, the prominence can rise into the corona and bow up into an archlike shape. A prominence can be thought of as a long magnetic tube whose length twists and turns much like a rope. Tight wrapping of the ropelike structure can give rise to various magnetic activities associated with phenomena such as flares or eruptive prominences.

Observations of sunspots made with the ultraviolet spectrograph provided information on the mass flow in the umbra (darkened central area) of sunspots and measures of the emission of radiation from them. With the ultraviolet spectrograph, scientists detected upward-propagating acoustic waves above the umbra, created by either a subphotospheric cavity trapping plasma waves or the resonant transmission of magnetoacoustic waves by a chromospheric cavity.

KNOWLEDGE GAINED

In the near UV, radiation with a wavelength of about 300 nanometers originates in the photosphere, and at progressively shorter wavelengths the radiation originates at higher and higher layers throughout the chromosphere. Shorter than 140 nanometers, the spectrum changes from a bright continuum with dark Fraunhofer lines to a faint continuum with emission lines. At wavelengths near 140 nanometers, the solar material is very opaque to radiation from the lower levels. As a result, it becomes possible to see emission lines from the chromosphere, since the continuum becomes so faint.

Many new emission lines were discovered by the Skylab and SMM spectrographs. The strongest emission line, the Lyman alpha line of hydrogen at 121.6 nanometers, is a very good solar activity indicator much like the line of ionized calcium at 392 nanometers. Images at the wavelengths of both lines show a mottled structure that fills in and becomes patchier and more irregular near centers of magnetic activity. Study of the Sun in various emission lines allows scientists to study higher and higher layers of the chromosphere and transition region. The Lyman alpha emission line, the strongest line in the entire solar spectrum, emits more energy than the entire solar spectrum between 0 and 120 nanometers.

The total solar UV irradiance is very low, accounting for about 1 percent of total solar electromagnetic emission. There is some evidence that the entire UV irradiance varies with the solar cycle. This variation has been established for EUV but not for all of the UV range. EUV irradiance is very sensitive to solar activity.

Radiation wavelengths of 100 nanometers or less have important terrestrial effects. This radiation originates in the chromosphere, transition region, and lower corona. The elaborate network structure of the chromosphere was verified and understood in much more detail than had previously been known from ground-based observations.

The transition region is striking in UV images. Since the temperature rises with height quickly in the transition region, the images taken in the EUV band show a pronounced limb brightening. The edges of these images are similar to a lightened ring surrounding the disk.

UV and EUV images of prominences show them to be coronal refrigerators. They are essentially tunnel-shaped magnetic tubes twisted like a rope, often into an arch rising above the chromosphere into the corona. The motion of material in the tube can be very complex. Motion occurs predominantly downward near the footpoints of the arches, whereas material near the top of the arch appears to have a buoyant quality. The arches are often seen to rise and stretch simultaneously until dissipating into the corona.

The ultraviolet spectrograph on the SMM verified that material motion, which had been detected in ground-based observations, occurs in the umbrae of sunspots. Ground-based observations suggest a motion starting at the center of the spot and radiating outward into a circle. In some cases, the reverse of that phenomenon—radiating motion inward to the center of the circle—was observed. The motions observed by the ultraviolet spectrograph seemed to be essentially vertical, rising outward from the spot. The motions also seemed to have a wavelike, pulsing character, suggesting a form of solar oscillations.

CONTEXT

Solar UV striking the upper atmosphere dissociates diatomic oxygen molecules (O_2) into separate oxygen atoms. Between 15 and 50 kilometers above Earth's surface, ozone (O_3) is formed by a combination of single oxygen atoms and diatomic oxygen molecules. Ozone is destroyed by UV radiation, particularly between 210- and 310-nanometer wavelengths. This radiation is harmful to living tissue. Equilibrium is normally maintained between the creation and the destruction of ozone. Earth's ozone concentration peaks at an altitude of about 20 kilometers, the "ozone layer," where its concentration is about ten parts per million. Nevertheless, this thin concentration is essential for the preservation of human life.

Knowledge of the UV and the EUV has allowed astronomers and solar physicists to study the elusive transition region between the chromosphere and corona. The transition layers glow mostly in the UV and show a pronounced increase in temperature and decrease in density with height in the solar atmosphere. Within these layers, various waves (both mechanically and magnetically generated) dissipate energy, mostly in the form of mechanical heat (the vibration of atoms and particles) to the layers above. Thus, they account for the steep increase in temperature with height in the chromosphere and corona above the photosphere. The pronounced decrease in density within these layers is partly responsible for allowing only certain types of wave energy to move outward, retaining most of the thermal energy. Various heat mechanisms, such as sound waves, induce vibrations among the particles, which in turn introduce heat and other forms of energy to their surroundings.

Most solar physicists still consider the increase in temperature of the chromosphere and corona with height above the photosphere to be one of the most difficult problems in solar physics. The foregoing explanations account for some of the mechanisms responsible, but not all. UV and EUV observations show fine structural details that may also be observable in X-ray wavelengths. These details may reveal phenomena that might contribute to this heating and also provide clues as to the physical nature of solar flares. Most notably, these details reveal the interaction of many small magnetic loops giving rise to new local magnetic geometries having the ability to accelerate charged particles and waves into the chromosphere and corona. Solar oscillations of various modes and amplitudes, the "ringing" of the Sun three-dimensionally, no doubt send energy to the chromosphere and corona.

James C. LoPresto

FURTHER READING

Bhatnagar, Arvind, and William Livingston. *Fundamentals of Solar Astronomy*. Hackensack, N.J.: World Scientific, 2005. An intermediary work between basic and advanced textbooks. In addition to solar astronomy, the author also provides directions for building a solar telescope and details how to observe the Sun at various portions of the electromagnetic spectrum.

Buchler, J. Robert, and Henry Kandrop, eds. *Astrophysical Turbulence and Convection*. New York: New York Academy of Sciences, 2000. Approaches the subject of turbulence and convection from four different perspectives: those of theorists, experimentalists, astrophysicists, and computational physicists. Technical, for scientists and college students studying the Sun and its energy production and emission.

Eddy, John A. *A New Sun: The Solar Results from Skylab*. NASA SP-402. Washington, D.C.: Government Printing Office, 1979. This spectacular volume highlights what was learned about the Sun as a result of the Skylab mission, using the Apollo Telescope Mount. Using spectacular photographs, the book contrasts astronomers' knowledge of the Sun before and after Skylab and shows how space-acquired information corrected and greatly amplified previous knowledge.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Provides discussions of solar energy production, output, and the influence it has on Earth's magnetosphere and environment.

Gibson, Edward G. *The Quiet Sun*. Washington, D.C.: Government Printing Office, 1973. This excellent survey of solar physics uses a structural outline form to discuss what is known about the Sun. The author starts with the solar interior and works his way outward through the atmosphere, discussing the photosphere, chromosphere, and corona in separate chapters. The first chapter is a very helpful overview of the entire topic. Solar activity and solar-terrestrial relationships are covered in the latter chapters. Rather technical; recommended for the reader with some background knowledge of the subject.

Golub, Leon, and Jay M. Pasachoff. *Nearest Star: The Surprising Science of Our Sun*. Cambridge, Mass.: Harvard University Press, 2001. The authors explain the contemporary state of knowledge about the Sun, including historical and observational data. An excellent source for nonscientists.

Haigh, Joanna, et al. *The Sun, Solar Analogs, and the Climate*. New York: Springer, 2005. A set of lectures published by the Swiss Society for Astrophysics and Astronomy. Focuses on how the Sun changes and how Earth's climate responds. Changes in solar ultraviolet output can have dramatic effects on the biosphere. For undergraduate and graduate students.

Hille, Steele, and Michael Carlowicz. *The Sun*. New York: Harry N. Abrams, 2006. A vast collection of images taken by photographers, satellites, and observatories.

Nicolson, Iain. *The Sun*. New York: Rand McNally, 1982. In an atlas-style presentation, Nicolson provides an overview of what is known about the Sun. Includes a historical perspective; discusses the relationship of the Sun to the stars, the Galaxy, and the universe; supplies a detailed description of the solar interior, the solar atmosphere, solar activity, and solar-terrestrial relationships. Many problems are discussed in detail, including solar flares, the solar neutrino problem, coronal holes, solar oscillations, and the solar wind. Even the pragmatic problems of solar energy are discussed.

Noyes, Robert W. *The Sun: Our Star*. Cambridge, Mass.: Harvard University Press, 1982. This excellent book reviews the research done in solar physics from ground-based telescopes and space satellites. The material presented is related to how modern knowledge of the Sun increases knowledge of stars in general. There is an emphasis on how studies of the solar spectrum provide information about the complex solar atmosphere. The contents are arranged according to regions or zones within the interior and atmosphere of the Sun. Much attention is given to solar activity and its relationship to Earth. Suitable for general audiences.

Stix, Michael. *The Sun*. New York: Springer, 2004. Covers all aspects of solar physics. Assumes the reader has some mathematics and physics knowledge. Includes practice problems on most topics.

Zirker, Jack B. *Total Eclipses of the Sun*. Expanded ed. Princeton, N.J.: Princeton University Press, 1995. A delightful short study

that discusses the history of observed eclipses as well as their significance to solar science. Current topics, such as using the Sun to test the general theory of relativity, the physics of the solar corona, and solar oscillations, are covered and explained well. There is also some discussion of the Sun's influence on Earth's atmosphere. An accessible presentation.

See also: Coronal Holes and Coronal Mass Ejections; Electromagnetic Radiation: Thermal Emissions; Interplanetary Environment; Nuclear Synthesis in Stars; Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Variability; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Variability

Category: The Sun

There is evidence that the Sun's total energy output varies. The variations are much smaller than the brightness variations observed in other types of variable stars, but even small changes in the Sun's energy output could have significant effects on Earth's climate. Variations correlated with the eleven-year sunspot cycle have been reliably measured. The evidence for possible longer timescale variations is not as clear.

OVERVIEW

A significant fraction of the stars we observe are "variable stars," which means that they vary in brightness. Many of these variable stars vary by significant fractions of their total energy output. Fortunately, the Sun does not fall into this class. If the Sun's energy output were to vary significantly, Earth's climate would be very unstable. However, if the Sun were to vary a very small amount, there might be small but measurable effects on Earth's climate.

The amount of solar energy falling on a square meter at the top of Earth's atmosphere every second is called the solar constant. Accurate measurements of the solar constant can reveal whether the Sun's total energy output per second, its luminosity, varies or is constant. However correcting for Earth's atmospheric absorption is difficult, making it difficult to measure the solar constant from Earth's surface accurately.

Samuel Pierpont Langley invented an instrument called a bolometer in 1878. A bolometer is an extremely sensitive electrical thermometer that can be used to measure the amount of radiant energy from the Sun at all wavelengths. Using his bolometer, Langley, who was secretary of the Smithsonian Institution, pioneered efforts to measure the solar constant accurately. To minimize possible errors from atmospheric absorption, Langley traveled to the top of Mount Whitney in 1881 to make his measurements. In 1895 Langley offered Charles Greeley Abbot a job at the Smithsonian. Using an improved bolometer, Abbot again measured the solar constant. Abbot spent the rest of his career accurately measuring the solar constant and looking for solar variability. Under his direction, the Smithsonian repeatedly measured the solar constant from 1902 to 1957. This program remains the longest continuous search for solar variability. Abbot and his assistants made observations from various remote high-altitude sites in his quest for accuracy. Abbot maintained that he had observed variability in the solar constant, but few other scientists thought that his data were accurate enough to detect solar variability.

Sunspots are dark areas on the Sun's visible surface, where solar magnetic fields divert energy coming from the Sun's interior. There is an eleven-year cycle in the number of sunspots on the Sun. From the perspective of modern satellite observations, it seems that solar variability is too small for Abbot's instruments to have detected. However, Abbot did turn out to be correct in his conclusions that when a major sunspot group faces Earth directly, the Sun's luminosity is slightly reduced and that paradoxically the Sun has a slightly higher energy output during the maximum sunspot activity. The Sun can have a greater luminosity during periods of maximum sunspots because during these times the number of faculae, which are brighter areas on the Sun, increase. Most scientists still thought, however, that it was just too difficult to correct for the variable absorption from Earth's atmosphere, and the question of solar variability remained controversial.

In 1978, the National Aeronautics and Space Administration (NASA) launched the Nimbus 7 satellite with a new, accurate radiometer for measuring the Sun's energy output. In 1980, the Solar Maximum Mission satellite also contained an accurate radiometer. These and later similar satellites confirm two of Abbot's conclusions. When a large sunspot group faces Earth directly for a short time, the Sun's irradiance decreases a small amount because the sunspots block some of the Sun's radiant energy. Over decade-long timescales, however, the Sun's total luminosity is about 0.15 percent higher during sunspot maximum than during sunspot minimum. This result is just slightly less than the estimated accuracy limit of Abbot's observations. During maximum sunspot activity the effect of the brighter faculae is slightly greater than the effect of the fainter sunspots, and the Sun's energy output is slightly higher.

In 1894, E. Walter Maunder called attention to the fact that very few sunspots were observed from 1645 to 1715. Few astronomers paid much attention to Maunder's work until the mid-1970's. In 1976 John A. Eddy published an article confirming this seventy-year paucity of sunspots and naming it the Maunder minimum. During this time, the Sun had about one-thousandth the normal number of total sunspots. Galileo did not discover sunspots until the early 1600's, so modern scientists use indirect means to deduce the number of sunspots prior to the time of Galileo. These indirect methods show the Spörer minimum about 1500, another minimum around 1350, and a medieval grand maximum in sunspot activity around 1200.

If, as satellite data show, the Sun's luminosity varies with the eleven-year sunspot cycle, it is reasonable to ask if it also varies with these much longer cycles. Because these longer cycles have greater variations in sunspot activity than the eleven-year cycle, it is also reasonable to ask if the Sun's luminosity variations are larger than during the eleven-year sunspot cycle. If they exist, increases or decreases in the Sun's total energy output lasting for several decades or longer should affect Earth's climate. If the Sun is less luminous during extended sunspot minima and more luminous during sunspot grand maxima, then the Maunder and Spörer minima should be periods of lower solar luminosity and the grand maximum should be periods of higher solar luminosity. Are they?

It is not possible to travel back in time to measure the Sun's energy output during these periods. However, extended periods of variable solar luminosity should affect Earth's climate. We also lack accurate climate records this long ago. It is therefore necessary to rely on anecdotal and indirect evidence to determine the Earth's climate, and by extension the Sun's luminosity, during these minima and maxima.

Direct measurements are not available, but there is considerable indirect evidence that the Sun's luminosity has varied over the last millennium and that these variations have affected Earth's climate. The time of the Maunder minimum is often referred to as the Little Ice Age. Considerable anecdotal evidence suggests that Europe was much colder than normal during this period. The indirect evidence includes reports of the Thames River freezing in England, Dutch paintings of frozen landscapes in places that usually do not freeze, tree ring studies, and other biological indicators. The grand maximum corresponds to the time period when a Viking colony flourished on Greenland, which is normally much colder than it was during the grand maximum. It is plausible that the Sun had greater luminosity during the grand maximum, and Earth's climate was warmer.

KNOWLEDGE GAINED

Descendants of Langley's bolometer are still in use today, although they are much more sensitive than Langley's original instrument. Modern astronomers cool them with liquid helium to increase their sensitivity and mount them on the ends of large research telescopes to measure the infrared energy from distant stars and galaxies. They have added considerably to our knowledge of the universe.

Abbot and his coworkers at the Smithsonian spent fifty years systematically measuring the solar constant from a variety of locations on Earth as accurately as possible. (It helped that Abbot lived to be 101.) Because solar variability is so small, he was not able to convince many scientists that he was really measuring the solar variability that he claimed. However, even scientists who disagreed with Abbot's claim that the solar constant varied recognized the precision and long time span of Abbot's data. Many scientists have since analyzed Abbot's data to study changes in the absorption of Earth's atmosphere, such as might be induced by major volcanic eruptions and similar phenomena.

The understanding of solar variability made a great leap forward when astronomers were able to put radiometers on satellites to measure the solar constant from space. Because astronomers did not have to correct for the changing effects of light absorbed by Earth's atmosphere, the resulting measurements of the solar constant were more accurate. Scientists could finally agree that the variability in the solar constant is real rather than a measurement error. As a result, efforts to understand how solar variability may be associated with climate changes over past millennia will result in a better understanding of Earth's climate, including current climate changes.

Just as solar variability may affect Earth's climate, it can also affect the temperatures on other planets in the solar system. For example, the polar ice caps on Mars could grow or shrink in response to long-term increases or decreases in the Sun's energy output. Earth, however, is the only planet in the solar system known to have life, so such temperature changes are most serious on Earth. If one of the other planets or satellites in the solar system does indeed have life, then this life could be adversely affected by solar variability.

CONTEXT

Global warming is considered to be among the most serious environmental issues facing the world today. Effects include rising sea levels, melting glaciers, ecosystem damage or destruction, increasing frequency of violent storms, and, in some regions, more severe droughts. Many but not all scientists agree that the major cause of this global warming is the increase in carbon dioxide and other greenhouse gases in our atmosphere caused by human activity. However, the causes may be more complex. If the Sun's luminosity does indeed vary over time periods of decades or centuries, these variations can also affect Earth's climate. To be accurate, climate prediction models need to take into account any possible solar variability as well as changes in the amount of greenhouse gases. If the Sun's luminosity is decreasing slightly, that could help solve the global warming problem. If the luminosity is increasing, however, that could exacerbate the effects humans have on Earth's climate. Unfortunately, at this point, solar variability is still too poorly understood to answer these questions definitively.

Paul A. Heckert

FURTHER READING

Eddy, John A. "The Maunder Minimum." *Science* 192 (1976): 1189-1192. In this seminal article, Eddy demonstrates that the Maunder minimum is real and argues that the Sun's energy output varies with long-term cycles in the amount of sunspot activity.

Frazier, Kendrick. *Our Turbulent Sun*. Englewood Cliffs, N.J.: Prentice-Hall, 1980. A good, readable account of our knowledge of the Sun through the publication date. Early studies in solar variability are covered in considerable historic detail.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. Chapter 16 of this introductory astronomy textbook is a complete, readable overview of our knowledge of the Sun.

Golub, Leon, and Jay M. Pasachoff. *Nearest Star: The Surprising Science of Our Sun*. Cambridge, Mass.: Harvard University Press, 2001. This well-written book gives a detailed summary of our knowledge of the Sun and includes discussion of solar variability.

Hester, Jeff, et al. *Twenty-First Century Astronomy*. New York: W. W. Norton, 2007. Chapter 13 of this astronomy textbook covers the Sun.

Hoyt, Douglas V., and Kenneth H. Schatten. *The Role of the Sun in Climate Change*. Oxford, England: Oxford University Press, 1997. Well written and extensively documented, this book completely covers the topic of solar variability and possible associated climate changes. The authors present both sides of controversial issues and try to appraise the role of solar variability in climate change realistically.

Maunder, E. Walter. "A Prolonged Sunspot Minimum." *Knowledge* 17 (1894): 173-176.

_______. "The Prolonged Sunspot Minimum, 1645-1715." *Journal of the British Astronomical Society* 32 (1922): 140ff. These two papers are Maunder's original publications arguing that there was a prolonged minimum in sunspot activity in the late seventeenth century. Maunder argues that the minimum was real and not an effect of few people observing the Sun.

Soon, Willie Wei-Hock, and Steven H Yaskell. *The Maunder Minimum and the Variable Sun-Earth Connection*. Hackensack, N.J.: World Scientific, 2003. This well-documented book explores the Maunder minimum and other long-term periods thought to result from solar variability, as well as the role this variability plays in Earth's climate changes.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. Cambridge, England: Cambridge University Press, 2002. An extremely well-written introductory astronomy textbook. Chapter 12 is an overview of our knowledge of the Sun.

See also: Auroras; Coronal Holes and Coronal Mass Ejections; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; Nuclear Synthesis in Stars; Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar Ultraviolet Emissions; Solar Wind; Solar X-Ray Emissions; Sunspots; Thermonuclear Reactions in Stars; Ultraviolet Astronomy.

Solar Wind

Category: The Sun

The Sun emits streams of protons, electrons, and some heavier particles in all directions. Known as the solar wind, the outward flow of material in these streams comes from the outermost region of the Sun's atmosphere, the corona.

OVERVIEW

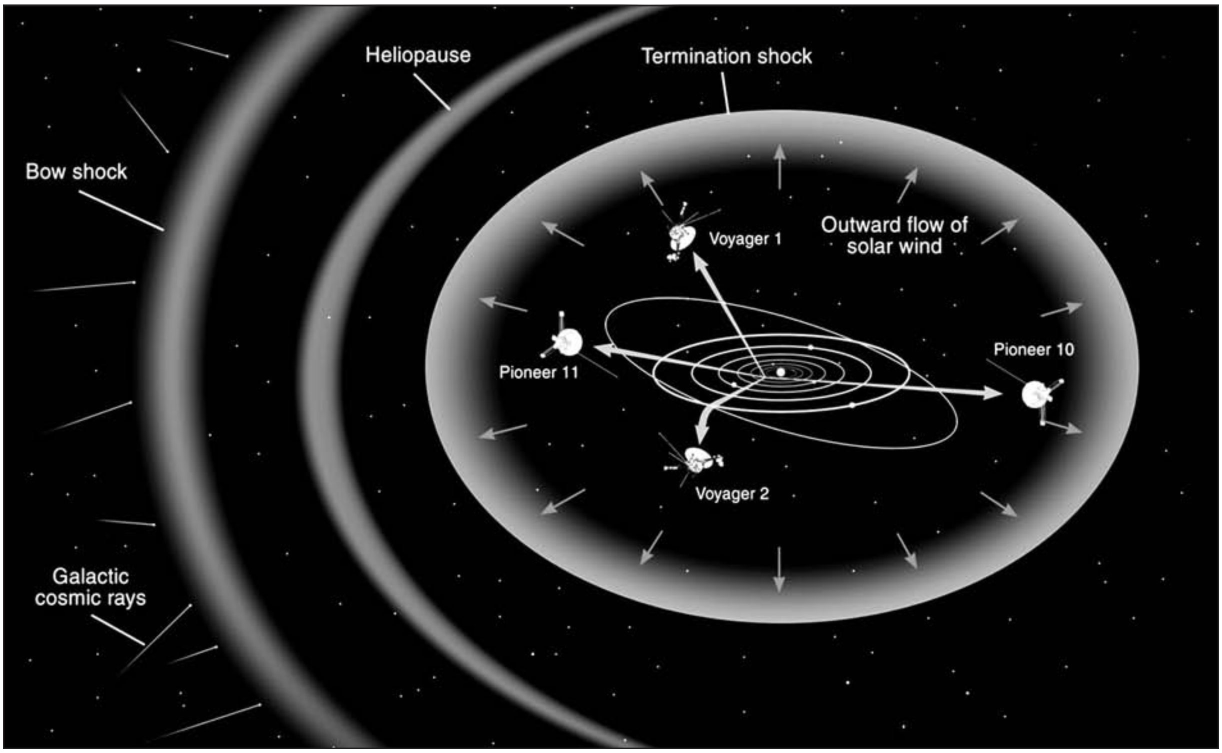
The Sun's corona (its outer atmosphere) does not end abruptly but gradually decreases in density, as it extends billions of kilometers into space. The outward movement and expansion of the corona are functions of distance from the Sun. Expansion close to the Sun is very slow, since the pull of gravity is dominant, but as the distance from the Sun increases, the outward flow increases. This flow of gas is the solar wind, the term originally devised by Eugene N. Parker in 1958 in his classic paper on the dynamics of the interplanetary gas. The solar wind is a stream of ionized gas constantly blown away from the Sun at high speed in all directions. It is composed primarily of electrons, protons, alpha particles (helium nuclei), and some heavier ions.

The rate of mass loss as a result of the outflow of material via the solar wind is only on the order of 10^{-14} or 10^{-15} of the solar mass per year (or about 10^{-9} of the Earth's mass per year). Since the Sun is losing mass, however, the total angular momentum of the Sun is decreasing. As a result, the angular speed of rotation of the Sun also decreases with time, at least on stellar evolutionary timescales. A study of surface rotation rates of young solar-type stars in the Pleiades and Hyades star clusters shows that they have rotation rates some ten times faster than that of the Sun.

At the distance of Earth's orbit from the Sun (150 million kilometers), the average number density of ions under "quiet Sun" conditions (periods when the Sun is not exhibiting high activity, as during solar maxima) is 5 particles per cubic centimeter. Varying solar activity can cause this number to vary widely from the average value; measurements from space probes yield a range of 0.4 to 80 particles per cubic centimeter. The temperature of the particles in the solar wind at the Sun is about 1 million kelvins. By the time they reach the Earth, their temperature has dropped to 200,000 kelvins on average, but because their density is so low, no appreciable heat transfer to Earth occurs. Again, there is considerable variation, from a minimum of 5,000 kelvins to a maximum of 1 million kelvins.

The use of the word "wind" is appropriate considering the speeds involved. At the Earth's orbit, the solar wind whips by at approximately 400 to 500 kilometers per second on average, though there are large fluctuations ranging between 200 kilometers per second minimum and 1,000 kilometers per second maximum. The solar wind is composed mostly of electrons and protons, with a helium abundance averaging 5 percent, but ranging from 0 percent up to a maximum of 25 percent.

Historically it was known that there is a correlation between solar activity and geomagnetic storms. A large summary of historical data correlating geomagnetic activity and solar activity was published in 1940 by Chapman and J. Bartels. In 1931, 9 years prior to this comprehensive summary, the initial model to provide an explanation for the connection was proposed by Sydney Chapman and V. C. A. Ferraro. The model involved streams of ionized (electrically charged) gas ejected from the Sun at the time of solar flares. The interaction of these ionized gas streams, trapped in the magnetic polar regions of the Earth, with the Earth's atmosphere triggered the northern and southern lights (also known as the aurora borealis and australis). This initial model was updated in 1951 by Ludwig Biermann, who suggested that, rather than occasional streams, there was a continuous outward flux of charged particles from the Sun into interplanetary space. This revised model was designed to explain the antisolar spikes observed in some comets. The classic paper by



A schematic showing the influence of the solar wind, the extent of which largely defines the heliosphere, which begins to slow dramatically at the heliopause and ends at the bow shock, the outer boundary of the solar system where the pressure of interstellar matter checks the solar wind. (Lunar and Planetary Institute)

Parker in 1958 predicted that interplanetary space is filled with a solar wind. The content and speed of this predicted wind would be functions of the temperature of the corona. Parker predicted that the solar wind speed in the vicinity of the Earth would range from 400 to 800 kilometers per second.

Parker's paper came out just at the dawn of the space age. Since then, numerous space missions have gathered massive amounts of data on the solar wind that could not have been obtained any other way. Between 1959 and 1961, the Soviet interplanetary probes Luna 2 and 3, Venus 1, and Mars 1, using rudimentary (by today's standards) particle detectors, confirmed Parker's prediction. Also in 1961, the initial findings of the Soviet space probes were confirmed by the United States' Explorer 10 satellite using a Faraday-cup probe. The first long-term monitoring of the solar wind was conducted by the U.S. Mariner 2 spacecraft on its three-month mission to Venus in 1962. Mariner 2 again confirmed Parker's prediction by measuring a continuous solar wind with speeds ranging from 319 to 771 kilometers per second. However, the Mariner results disclosed something not predicted by Parker: gusts in the solar wind up to 1,000 kilometers per second that were phase-correlated with the rotational period of the Sun.

The reason for the correlation between the Sun's rotation, solar wind gusts, and geomagnetic perturbations with a recurrent twenty-seven-day period (which matched the Sun's synodic rotation period) was not determined until rocket and spacecraft X-ray data of the Sun revealed the existence of nonmagnetic, long-lived holes in the corona. The coronal holes rotate with the rest of the Sun, and, since they are nonmagnetic, they permit ions an easy exit from what normally are magnetically confined regions of the Sun. The streams of ions passing through the coronal holes have higher-than-normal speeds because they are able to escape without having to overcome the additional magnetic force effects usually present at the Sun's surface. Data from Skylab firmly established that coronal holes are the source of the high-speed streams (gusts) of the solar wind. This discovery is among the most well-established solar-terrestrial connections and is useful in making daily forecasts of geomagnetic disturbances.

Even though the solar wind is emitted in all directions from the Sun, the resultant plasma is not uniform. Observations indicate that the speed of the solar wind tends to be higher and more variable at high solar latitudes. The polar regions of the Sun are nearly always covered by coronal holes, and thus one would expect gusts from the higher latitudes at higher-than-normal speeds.

The Pioneer series of spacecraft carried instrumentation to detect and measure the solar wind. In August, 1972, Pioneer 9 (at a distance from the Sun close to Earth's) recorded solar wind speeds of 1,000 kilometers per second, while Pioneer 10 (which was 214 million kilometers from the Sun, nearing the orbit of Mars) recorded the solar wind at about half that speed. In 1983 Pioneer 10 detected the presence of the solar wind as far out as 4.5 billion kilometers, at the orbit of Neptune. All the Pioneer data show that the average speed of the solar wind changes comparatively little out to Jupiter's orbit, but the range of fluctuations in speed is remarkably diminished at Jupiter's orbit compared with the range at Earth's orbit. This can be seen clearly in the Pioneer data.

KNOWLEDGE GAINED

The bulk of the solar wind is composed of electrons, protons (hydrogen nuclei), and alpha particles (helium nuclei), but there are also traces of heavier elements, the most abundant of which are measurable at the distance of Earth's orbit. Solar wind abundances have been derived from Vela (3, 5, and 6) and Apollo (11, 12, 14, 15, 16, and 17) mission data for the elements that are most common in the entire solar system and in the Sun's corona. The agreement of relative abundance values is quite good considering the varying accuracies of the individual determinations (roughly a factor of two).

The Apollo 17 mission provided data to compute the abundance of iron in the solar wind. Other Apollo missions produced data that resulted in abundance ratios for the light noble gases neon and argon and their isotopes. A major finding of the Apollo program was the detection of the heavier noble gases krypton and xenon in the solar wind and, by implication, the presence of these two elements in the Sun. The anomalous presence of noble gases in meteoritic and lunar surface material is explained by the exposure of these materials to the solar wind.

The ancient solar wind is preserved in approximately the outermost micron of the surface of solid objects, since ions with kinetic energies greater than 1 kilo-electron volt (keV) are trapped on impact with solid surfaces. This record is an easy target for alteration or destruction by a whole host of events. Even so, analyses of lunar material and meteorites indicate that some sort of solar wind has been present at the distance of Earth's orbit for the past 3 to 5 billion years. A study of Apollo 15 deep-drill cores (as deep as 3 meters) by Donald D. Bogard and L. E. Nyquist concluded that the solar wind at the distance of Earth's orbit shows little variation over the last 400 million years.

Magnetic fields associated with the planets in the solar system are distorted by interaction with the solar wind. The magnetopause (outer boundary of a planet's magnetosphere) is where a balance exists between the solar wind and the planet's magnetic field. The magnetopause boundary appears to be stable for planets with large magnetic fields. An average value for the size of Earth's magnetosphere (a term in which the root "sphere" does not necessarily mean spherical in shape, but rather refers to the sphere of influence of Earth's magnetic field) is about 13 Earth radii. This is a sizable obstruction in the path of the solar wind. The result of this obstacle to the supersonic flow of the solar wind is a standing shock wave. In order to have a smooth flow about the periphery of the magnetosphere, the solar wind must make the transition from supersonic to subsonic speeds. In addition, when the plasma is so abruptly slowed, its effective temperature is increased (about tenfold). The point of closest approach by the solar wind to the center of a planet is called the stagnation point. The stagnation point for Earth is at a distance of about 10 Earth radii from the center of the planet (well above the atmosphere).

On December 27, 1984, the Active Magnetospheric Particle Tracer Explorers (AMPTE), a three-satellite cooperative venture by West Germany, Britain, and the United States, generated an artificial cloud of ionized barium. The project was designed to track ionized elements to determine how many of the solar wind's ionized particles actually enter the Earth's magnetosphere and to understand the formation and motions of the high-energy particles in the Van Allen trapped-radiation belts. Results from an earlier test in September, 1984 (with lithium as the tracer element), had shown that approximately 1 percent of solar-wind particles are transported into the magnetosphere surrounding the Earth.

The data from these tests were not definitive and provided only a measure for the specific conditions existing at the time of the test. They indicated that charged particles captured from the solar wind spill out of the outer Van Allen radiation belt (which is inside Earth's magnetosphere) and enter Earth's atmosphere around the north and south magnetic poles. The collisions of these charged particles with atoms of oxygen and nitrogen stimulate them to radiate pale greens and bright reds in the northern and southern skies. This colorful display is known as the aurora borealis or aurora australis (the northern or southern lights), most often seen in zones between 65° and 70° north and south magnetic latitude.

The heliopause is the boundary about 15 billion kilometers from the Sun (about three times the size of Neptune's orbit), where the solar wind merges with the interstellar medium and loses its identity. The interface is the outer limit of the heliosphere, the region influenced by the solar wind. Pioneer 10 data showed that the solar wind oscillates with the eleven-year solar cycle. Increases in solar activity and the solar wind result in decreasing numbers of entering galactic cosmic rays, since the more active heliosphere acts as a shield. Thus, the solar wind influences both the solar and galactic cosmic-ray fluxes at the Earth through a modulation process at the heliopause and through magnetic field-line reconnections at the magnetopause.

CONTEXT

The nucleus of a comet is a "dirty snowball" of various ices with embedded dust and grit. The outer layer of the nucleus sublimates when it nears the Sun, producing a gaseous cloud (the coma) around the nucleus and one or two tails—Type I and Type II. Type II tails are the dust tails. They are smooth, homogeneous, and point generally away from the Sun along a curve. Dust grains in Type II comet tails have a high area-to-mass ratio and thus the effect of solar radiation pressure is significant; Type II tails are blown away from the Sun by the transfer of momentum from solar photons to the dust grains. Type I tails are the ionic tails and are primarily composed of CO^+ along with some other ions, such as N_2^+, CO_2^+, CH^+, and OH^+. Type I tails are long, straight, patchy, and point radially away from the Sun.

Early theories postulated some sort of interplanetary medium to account for the orientation of Type I ionic tails. An extensive study of Type I comet tail orientations by C. Hoffmeister in 1943, long before Parker's predictions and the actual discovery of the solar wind, required the existence of an interacting resistive medium expanding outward from the Sun. A detailed analysis by John C. Brandt yielded an average expansion speed of 474 ± 21 kilometers per second, with a minimum near 150 kilometers per second. These values match very well the speeds measured for the solar wind by spacecraft. The interaction of the solar wind with the coma of a comet produces a bow shock, a hundred thousand to a million kilometers from the nucleus. This interaction in turn carries charged particles from the coma away from the Sun, forming the Type I tail.

In 1981, Brandt and Malcolm B. Niedner published an extensive photographic summary of Type I tail disassociation events (when the ionic tail disconnects from the comet). An angular sector structure of the interplanetary medium and solar wind was identified that rotates with the Sun, within which the direction of the dominant magnetic field is constant. The basic premise is that as a comet crosses a sector boundary, the sudden reversal of the dominant magnetic field direction causes the charged tail to break away from the rest of the uncharged comet, producing the patchy appearance of Type I tails. Thus, comets can be used as interplanetary magnetic probes to determine the spatial location of the sectors.

Theresa A. Nagy

FURTHER READING

Akasofu, Syun-Ichi, and Y. Kamide, eds. *The Solar Wind and the Earth*. Boston: D. Reidel, 1987. An advanced text that deals with a variety of aspects of the solar wind as it interacts with the Earth. The introductory portion of each of the fourteen chapters is quite easy to understand, and the degree of difficulty varies as a function of the topic. A very useful source book.

Brandt, John C. *Introduction to the Solar Wind*. San Francisco: W. H. Freeman, 1970. A good introduction to the subject of the solar wind with easy-to-read chapters on the historical summary, ground and space observations, and the interaction of the solar wind in the solar system. A few chapters are for the more advanced technical reader. Somewhat dated.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Discusses the solar wind and its effects.

Foukal, Peter. *Solar Astrophysics*. 2d rev. ed. Weinheim, Germany: Wiley-VCH, 2004. New York: Wiley, 1990. A detailed look at the Sun and our understanding of it, including the solar wind. Also covers the history of solar astrophysics. Suitable for undergraduates.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Includes material on the solar wind and its impact on the planets and solar system.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook. Thorough and well-written. Covers the solar wind.

Gosling, J. T., and A. J. Hundhausen. "Waves in the Solar Wind." *Scientific American* 236 (March, 1977): 36-43. Variations of the solar wind are described. For a wide audience.

Meyer-Vernet, Nicole. *Basics of the Solar Wind*. Cambridge, England: Cambridge University Press, 2007. An introduction to solar wind. Covers the Sun's structure, interior, and atmosphere. Ideal for researchers and astronomy majors.

Moldwin, Mark. *An Introduction to Space Weather*. Cambridge, England: Cambridge University Press, 2008. A textbook designed for nonscience majors, giving a solid introduction to the Earth-Sun system, space physics, and space weather. Includes definitions of key terms, underlying physics concepts, and review questions.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Thorough introduction to astronomy, divided into short sections on specific topics, including the solar wind and its effects.

See also: Auroras; Coronal Holes and Coronal Mass Ejections; Earth-Sun Relations; Interplanetary Environment; Planetary Magnetospheres; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar X-Ray Emissions; Sunspots.

Solar X-Ray Emissions

Category: The Sun

The Sun emits X rays and gamma rays, revealing the presence of extremely high temperatures (tens of millions of kelvins) and high-energy particles. Solar X rays are produced in the solar corona (the hot, tenuous outer atmosphere of the Sun), and both X rays and gamma rays are generated by the magnetic explosions or eruptions known as solar flares.

OVERVIEW

X rays and gamma rays are forms of electromagnetic radiation with very short wavelengths and very high energy photons. Electromagnetic radiation displays the properties of both waves and particles; the term "photon" is used to refer to electromagnetic radiation when it acts like a stream of particles. The wavelengths of X and gamma rays are less than about 10 nanometers, compared to visible light, which ranges from about 400 nanometers (violet) to 700 nanometers (red). Their photons have energies ranging from about 100 electron-volts (eV) upward; for comparison, the photons of visible light have energies from a little less than 2 to a little more than 3 eV. Because their photons are so energetic, the presence of X rays and gamma rays generally indicates high temperatures and interactions of high-energy particles.

The existence of temperatures of at least a million kelvins in the solar corona (the Sun's outer atmosphere) had been known since the 1940's because of the presence of emission lines of highly ionized elements in its optical spectrum. Why should the corona be so hot? The photosphere, or visible "surface" of the Sun, is much cooler (around 5,800 kelvins), and common sense had misled astronomers to expect a steady decrease of temperature outward, rather than the precipitous temperature increase that actually occurs.

The physical cause of the high temperatures of the solar corona now is thought to be the strong magnetic fields in the Sun. The Sun's magnetic field is complex, with loops or arches that extend beyond the photosphere out into the corona. Charged particles flow along these loops, gain energy from them, and transfer it to the corona. Observations show that the corona is not uniform at X-ray wavelengths. The strongest, "brightest" X-ray emission comes from regions where the magnetic field is the strongest and the gas is hottest. Coronal streamers of hot ionized gas follow the Sun's magnetic field lines outward from these areas. Coronal holes are large, dark (in X rays), cool regions with weak or absent magnetic fields.

Solar flares were originally discovered serendipitously in 1859, in ordinary visible light, by Richard Carrington, who had been making routine observations of sunspots. Subsequent observations at certain specific wavelengths (such as the red light at 656.3 nanometers emitted and absorbed by hydrogen atoms) disclosed flares more distinctly in the solar chromosphere (that part of the Sun's atmosphere in between the photosphere and the corona), so before the 1950's the general phenomenon was known as "chromospheric flares." They were not assigned much importance physically.

X-ray and gamma-ray observations changed this perception of the importance of solar flares. Whereas solar flares had been perceived as complicated but otherwise insignificant features of the solar atmosphere—some kind of solar cloud—they were discovered to be a fundamentally important phenomenon, the prototype object of high-energy astrophysics. This branch, which involves observations of X rays and gamma rays, brings together gravitational, atomic, nuclear, and plasma physics.

X-ray bursts accompany essentially every solar flare, and during times of great solar activity, this occurs many times per day. The "soft" (longer-wavelength, lower-energy) X-ray spectrum reveals the existence of dense plasmas, hot ionized gases composed of electrons and atomic nuclei with temperatures in the range of 10 to 20 million kelvins, some ten times the typical coronal temperatures. X-ray images obtained with grazing-incidence telescopes show the hot plasmas to be trapped in the magnetic tubes, or loops, that may rise about 100,000 kilometers above the photosphere of the Sun. The energy trapped in the hot plasma flows down the magnetic field lines over a period of minutes, feeding into the chromosphere. The plasma, shown by its soft X-ray emission, appears to be the central agent in producing many of the classical effects of solar flares.

What initially produces the hot flare plasma? Its creation is usually marked by a "hard" X-ray burst (shorter-wavelength, higher-energy X rays). These hard X rays come from the interactions of fast electrons, with energies far above those of the particles in the hot plasma responsible for the soft X rays. Furthermore, data from the Solar Maximum Mission (SMM) have shown that gamma-ray bursts also occur commonly in solar flares. The presence of gamma rays indicates that the high energy required to produce flares involves the acceleration of protons (and other ions) as well as electrons. X rays and gamma rays have revealed many of the mechanisms involved in flare events. Magnetism plays a crucial but ill-understood role in causing the plasma instabilities that put on such spectacular and dramatic displays in solar flares. However, the initial cause of these events—and hence a general theory of flares—remains elusive.

To summarize what we do know, a solar flare is now believed to originate as an instability occurring in the solar atmosphere, a heterogeneous magnetized plasma that extends upward from the photosphere (visible surface of the Sun) into interplanetary space. This instability results in the acceleration of high-energy particles, both electrons and ions, and the creation of plasmas with temperatures of tens of millions of kelvins. In essence, a flare is a magnetic explosion in the upper solar atmosphere, leading to rapid acceleration of high-energy particles and an outward eruption of denser solar material from near the photosphere. The flare features observable by ordinary techniques of optical spectroscopy from ground-based telescopes appear to be secondary products of this explosive release of magnetic energy.

KNOWLEDGE GAINED

Solar flares have connections to Earth that are both economically significant and scientifically important. Their high-energy radiation can perturb Earth's ionosphere (a layer of ionized atoms in the upper atmosphere), inducing surge currents to flow in electrical power grids, causing failures and sometimes extensive blackouts. Similar disturbance of the ionosphere can also interrupt radio communications, because the ionosphere reflects some types of radio communication. Unfortunately, solar physicists cannot predict the level of solar activity with much more precision than is afforded by the simple recognition of the well-known eleven-year sunspot cycle. They can predict large flares no better than seismologists can predict major earthquakes. The problem appears to lie in the complexity of the physics and the lack of adequate observations.

The high-energy radiation of a solar flare is produced in various ways. Soft X rays are produced by hot plasmas (with temperatures exceeding 10 million kelvins). In a major solar flare, the X-ray flux may reach a level as high as one-millionth of the total solar luminosity. Hard X rays are intense flashes of higher-energy X rays that occur near the onset of a solar flare, showing the acceleration of energetic electrons. Gamma rays are produced by nuclear interactions in solar flares, showing that particle acceleration in solar flares extends to protons and other heavier ionized particles. The spectra, time profiles, and spatial distributions of the high-energy radiation all serve as guides to the physics of the flare phenomenon. Fairly detailed observations of some of the properties of soft X rays have been obtained, but many observational gaps exist for hard X rays and gamma rays.

The introduction of grazing-incidence X-ray optics, first from sounding rockets and later (in 1973) from the Skylab crewed space station, showed that the hot plasmas responsible for X-ray emission from solar flares were trapped in magnetic loops. These structures have their "footpoints" anchored in the solar photosphere, but extend great distances up into the corona. Numerous X-ray observatories since the Skylab era have been used to study solar flares and the Sun's corona such as XMM-Newton and the Chandra X-Ray Observatory.

CONTEXT

The discovery of X rays and gamma rays from the Sun led directly to a new branch of astronomy, high-energy astrophysics, and to great changes in understanding solar behavior. These discoveries resulted from the use of instruments designed to detect these high-energy emissions, mounted on instrument platforms ranging from the V-2 rockets captured from Germany during World War II to high-altitude uncrewed balloons and eventually artificial Earth satellites and deep space probes. These vehicles were able to place these detectors above Earth's atmosphere, which blocks high-energy photons, and opened a whole new view of the Sun and other stars.

After World War II, when American researchers made observations above Earth's atmosphere, initially using V-2 rockets captured from the Germans. These observations showed the somewhat unexpected existence of "high-energy" radiation from the Sun, from ultraviolet (UV) to X rays. Herbert Friedman's group at the Naval Research Laboratory was instrumental in these observations, which paralleled those of James Van Allen, the discoverer of Earth's radiation belts. A most important discovery came in 1958, when Laurence E. Peterson and John Winckler flew a high-altitude balloon over Cuba for cosmic-ray studies and were able to detect gamma radiation from a solar flare that happened to occur during the balloon flight.

Over the decades since the 1950's, new launch vehicles, combined with remarkable progress in the technology of X-ray optics and X-ray and gamma-ray detectors, have led to a great expansion of research in this area. Solar observations in the X-ray and gamma-ray spectral regions have greatly broadened our knowledge of the physics of solar flares and hence the Sun's structure and physics.

In addition, X rays and gamma rays are now observed from a variety of other astronomical sources: white dwarfs, neutron stars, black holes, supernova remnants, galaxies, and clusters of galaxies, to name some. Observatories specifically designed to detect the X rays and gamma rays of such objects, including the Chandra X-Ray Observatory, the Compton Gamma Ray Observatory, XMM-Newton, the Swift mission, and Astro-E2, have been launched above Earth's atmosphere and have returned data that have rendered startlingly beautiful images of some of these sources. It can be speculated that some of the same physics of magnetized plasmas in the Sun may underlie the high-energy emissions from these objects.

Hugh S. Hudson

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well written college-level textbook for introductory astronomy courses, with material on X-ray observations, solar flares, and the Sun's corona.

Eddy, John A. *A New Sun: The Solar Results from Skylab*. NASA SP-402. Washington, D.C.: Government Printing Office, 1979. The focus here is on the abundant illustrations, mostly photographs obtained from the Skylab satellite, supplemented by discussion of the nature of the Sun. Includes some treatment of the phenomenon of solar flares and their radiation.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. This well-written college textbook for introductory astronomy courses covers X-ray observations of solar flares and the Sun's corona.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook, thorough and well written. Has sections on X-ray observations, solar flares, and the Sun's corona.

Giovanelli, Ronald. *Secrets of the Sun*. Cambridge, England: Cambridge University Press, 1984. A general introductory description of the Sun and of solar physics.

Noyes, Robert W. *The Sun: Our Star*. Cambridge, Mass.: Harvard University Press, 1982. Provides a general survey of knowledge of the Sun, with some material on processes leading to X-ray emission. This book is at a slightly higher technical level than the others in this bibliography, but is still suitable for a general readership.

Ripley, S. Dillon. *Fire of Life: The Smithsonian Book of the Sun*. Washington, D.C.: Smithsonian Exhibition Books, 1981. A beautiful, lavishly illustrated popular description of the Sun. Includes a discussion of societal impacts of solar phenomena.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses. Divided into lots of short sections on specific topics, with several touching on X-ray observations, solar flares, and the Sun's corona.

See also: Coronal Holes and Coronal Mass Ejections; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; Red Giant Stars; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar System: Origins; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Sunspots; Thermonuclear Reactions in Stars; X-Ray and Gamma-Ray Astronomy.

Space-Time: Distortion by Gravity

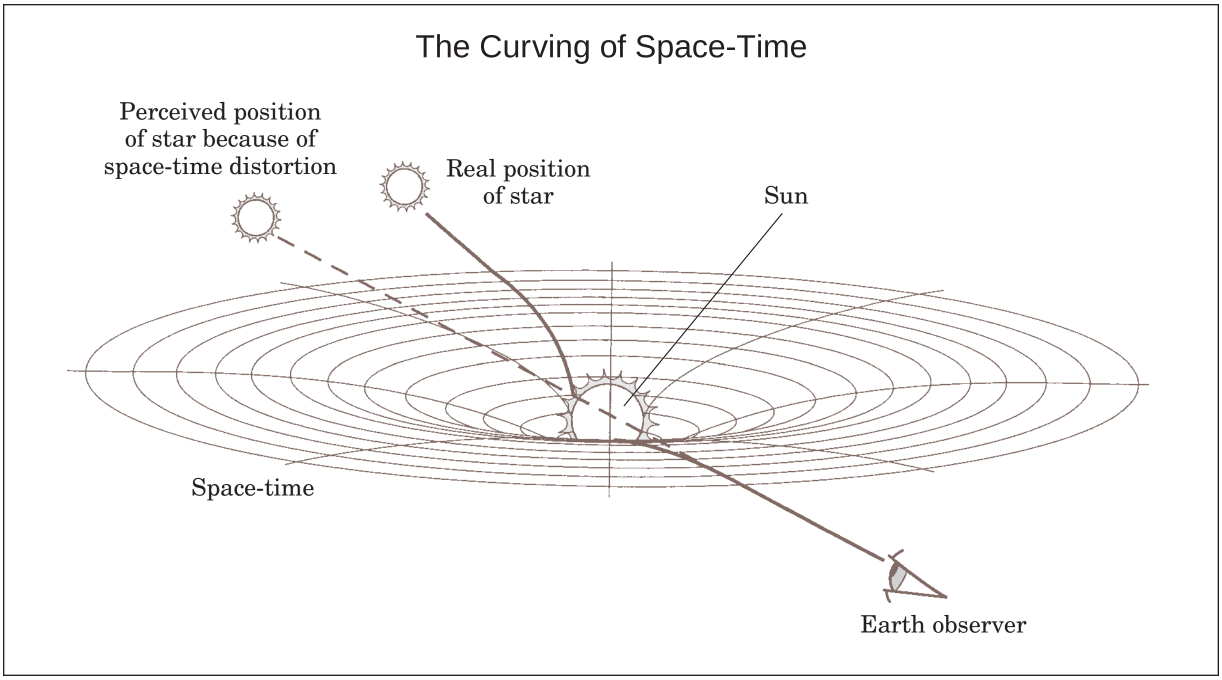
Category: The Cosmological Context

Space and time are linked together, and the fabric of space-time can be distorted or warped by the presence of matter. Albert Einstein's general theory of relativity reinterprets many diverse phenomena by replacing Newtonian gravity with distortions of space-time.

OVERVIEW

To the layperson, space is empty but time is full of activity. Time is perceived as a flow, carrying consciousness from one present moment to the next. To a physicist, however, the terms "space" and "time" denote quite different concepts: Space possesses physical properties and many levels of structure, and the flow of time is an illusion. Indeed, some theories hold that matter, rather than being located in space and time, is nothing more than disturbances in the fabric of space and time.

Although space and time can be perceived in radically different, separate ways, they are linked together by motion: Average speed is defined as distance divided by time. The motion of material objects and light signals is described in relativity theory in terms of a single, unified reality called space-time. Anything that happens in some particular place at some particular time is called an "event" in space-time. Its location in space-time is specified by its three spatial coordinates and one time coordinate. The series of events that trace the "where" and "when" of an object is called its world line through space-time. Objects at rest or moving with a constant velocity follow straight world lines, while the world lines of accelerated objects are curved.



The Sun's gravity causes space-time to curve, which bends the star's light and makes it appear to be located where it is not.

The structure of space can be described mathematically in terms of geometries. The geometry taught in most high schools is called Euclidean geometry. Developed (or at least described) by Euclid around 300 B.C.E., it deals with "flat" space. Some familiar features of flat space are that parallel lines never meet, even when extended infinitely far in both directions, and that the sum of the interior angles of a triangle equals 180°. During the mid-nineteenth century, the Russian mathematician Nikolai Ivanovich Lobachevsky (1793-1856) and the German mathematician Georg Friedrich Riemann (1826-1866) developed self-consistent non-Euclidean geometries. In these new geometries, the sum of the angles of a triangle could be more or less than 180°, and locally parallel lines could converge or diverge over large distances. A physical space described by a non-Euclidean geometry is called a curved space, as opposed to flat Euclidean space.

A familiar curved two-dimensional (2-D) space is the surface of a sphere, such as the Earth or a basketball. If one draws a triangle on the surface of a sphere, and that triangle is small (compared to the radius of the sphere), the angles of the triangle add up to 180° because the surface is nearly flat over small distances. However, if the triangle covers a large part of the surface, the angles add up to more than 180°. As an example, imagine doing this on the surface of the Earth. One side of the triangle is the 0° longitude line between the North Pole and the equator. The second side of the triangle is the 90° longitude line, also between the North Pole and the equator. The third side of the triangle is the equator from 0 to 90° longitude. Each of the three angles of this triangle is 90°, and thus the sum of the angles is 270°. Similarly, lines of longitude are parallel where they cross the equator, but because they are located on a sphere, they eventually meet at both poles.

These results—the angles of a triangle adding up to more than 180°, and locally parallel lines converging—are characteristic of, and can be used as tests for, the geometry developed by Riemann for positively curved spaces. The opposite properties—the angles of a triangle adding up to less than 180°, and locally parallel lines diverging—are characteristic of, and can be used as tests for, the geometry developed by Lobachevsky for negatively curved spaces.

The mathematics of curved space geometries took on new physical significance in 1915, when Albert Einstein (1879-1955) put forth his general theory of relativity. One consequence of it is that four-dimensional (4-D) space-time, first introduced by Einstein in his 1905 special theory of relativity, is curved by the presence of matter. Gravity appears to exist only because mass warps space-time in its vicinity. Physicist John Wheeler, who contributed enormously to relativity theory, provided this succinct description of the relation between matter and space-time: Matter tells space-time how to curve, and the curvature of space-time tells matter and light how to move. Where space-time is not flat, Euclidean geometry fails to provide an adequate description; non-Euclidean geometries must be used instead.

The fundamental postulate at the heart of general relativity is the principle of equivalence: The effects of a uniform gravitational field and the effects of a constant acceleration are indistinguishable. In a closed rocket accelerating through space at 9.8 meters per second per second (m/s^2), there is no experiment that an astronaut could perform that would distinguish this situation from one in which the rocket is at rest on the Earth's surface (where the acceleration of gravity is 9.8 m/s^2). Einstein realized that, if he could express this equivalence in a mathematical form, then he could relate both gravitation and acceleration to the curvature of space-time, thereby providing a geometric explanation of gravity. Non-Euclidean geometry provided the mathematical formalism to describe how the distribution of mass warps space-time. Einstein's field equations of general relativity enable one to compute, in principle, the curvature or warpage in space and time due to a given mass distribution.

Naturally, the bending of space-time profoundly affects the world lines of objects. As space-time warps, the world lines warp as well, since they are constrained to follow geodesics (locally straight lines) through a curved space. When an object moves on a world line that is curved, the object is being accelerated; that is, its motion is not uniform. From the perspective of general relativity, gravity is not to be considered as a force, but rather as an effect of geometry—a distortion of space-time. The presence of mass will distort the fabric of space-time. What is experienced as gravitation is the warping of space-time because of mass.

APPLICATIONS

The idea that matter distorts space-time led Einstein to make several remarkable predictions. In Euclidean geometry, the shortest distance between two points is a straight line. In a curved space, the shortest path is curved because space itself is curved. In order to gain a clearer understanding of this concept, consider that New York and Tokyo have similar latitudes but that the shortest distance between them, along the Earth's surface, passes close to the North Pole. Although incomprehensible on a flat map, this fact is easily visualized on the curved surface of a globe.

Similarly, a light beam passing near a massive object will follow a curved path because the massive object distorts space-time in its vicinity. The bending of light in the presence of a massive object also is mandated by the equivalence principle. In a reference frame accelerating upward, a transverse light beam would appear to deflect toward the upwardly accelerating floor. Since the equivalence principle requires that there be no distinction between acceleration and gravitational fields, light has to be deflected toward a source of gravity—that is, a massive object that warps space-time around it.

In 1911, Einstein predicted that a ray of starlight passing near the Sun would be bent by the warpage of space-time around the Sun. (Stars can be seen near the Sun during total solar eclipses when the sky is darkened sufficiently.) Although the predicted deflection of starlight just grazing the Sun's limb (edge) is only 1.75 arc seconds, it was first confirmed at the 1919 total solar eclipse, and later at many subsequent total solar eclipses since then. The effect has also been verified at radio wavelengths when the Sun occults (passes in front of) the radio source 3C279 every October 9. A similar effect, called gravitational lensing, has been observed when light from very distant galaxies is bent when it passes near massive, closer galaxies.

A previously known but unsatisfactorily explained effect—the advance of Mercury's perihelion—also is explained by the warpage of space-time near the Sun. Mercury's orbit around the Sun is slowly precessing (or pivoting) so that Mercury reaches perihelion (the point when it is closest to the Sun) slightly later with each orbit. Observations of Mercury's motion were sufficiently accurate that its perihelion advance was know by the mid-nineteenth century. Almost the entire effect could be explained by the gravitational perturbations of the other known planets on Mercury, but a small residual was attributed to the gravitational perturbation of an as yet undiscovered planet (given the name Vulcan) inside the orbit of Mercury. Vulcan was searched for, and some claims of discovery even were made, though none was confirmed. In 1915, Einstein explained how the warpage of space-time around the Sun causes Mercury's orbit to pivot forward (or precess); the precession predicted by general relativity accounts almost precisely for the residual that once was attributed to Vulcan. This (along with the bending of starlight passing near the Sun) was one of the first predictions of general relativity to be confirmed observationally. Since then, residual perihelion advances have been measured for Venus, Earth, and the asteroid Icarus; all agree within their observational errors with the predicted values calculated from general relativity.

Since space and time are intrinsically bound together, the effect of mass on the geometry of space-time means that time will be warped as well as space. Einstein predicted that time would be slowed by the warped space-time in the vicinity of a mass. Thus, clocks at the Earth's surface should run slightly slower than clocks at higher altitudes, which are farther from Earth's mass, where the warpage is less. The effect is small (about one part in 10^{13} for each vertical kilometer of altitude), but this minuscule effect was first measured (indirectly) in 1960 by the physicists R. V. Pound and G. A. Rebka at Harvard, again verifying a prediction of general relativity.

CONTEXT

In 1827, the German mathematician Carl Friedrich Gauss published a paper in which he recorded his measurements of the interior angles of a large triangle formed by three mountain peaks. His measurement was an attempt to ascertain whether space was Euclidean. The experiment was inconclusive: He obtained 180° within the experimental accuracy of his measuring devices. Gauss's experiment may have been the first attempt to test the long-established assumption that the universe is best described by Euclidean geometry.

By the middle of the nineteenth century, two other self-consistent geometries had been devised: the geometry of negatively curved space, by the Russian mathematician Nikolai Ivanovich Lobachevsky, and the geometry of positively curved space, by the German mathematician Georg Friedrich Riemann. Until 1915, however, when Einstein published his general theory of relativity, most mathematicians and physicists assumed that non-Euclidean geometries were mathematical curiosities but had little to do with physical reality. In Einstein's theory, space-time is curved by the presence of mass, and gravity exists only because mass gives space a non-Euclidean character.

Several months after Einstein's theory was published, the German astrophysicist Karl Schwarzschild found rigorously exact solutions to Einstein's field equations for two different cases: an ideal point mass and a finite spherical mass. The first case predicted that, at a relatively small radius from the mass point, some of the mathematical terms become infinite. This condition represents such an intense warping of space-time that any signal (whether matter or electromagnetic radiation) within this boundary would be unable to escape. This radius, called the Schwarzschild radius, defines a surface (called the event horizon) such that any mass or energy within this surface is forever trapped. Most physicists of the time, Einstein included, believed that it would be impossible for any real object to contract to such a small size that its mass would be contained within this surface. (For example, an object with the mass of the Sun would have to shrink within a radius of about 3 kilometers, and an object with the mass of the Earth would have to shrink within a radius of about 9 millimeters.) In 1963, Roy Kerr developed a new exact solution to Einstein's equations for a rotating mass, and the Schwarzschild solutions were seen to be special cases of Kerr's solution. Since matter and energy could cross the event horizon in the inward direction but nothing could escape from inside the event horizon, the term "black hole" was coined.

By the 1970's, indirect evidence had begun to accumulate indicating the existence of just such black holes, and there now is compelling evidence for at least two classes of black holes. Stellar-mass black holes form when massive stars exhaust all the fuels they used for nuclear fusion reactions and finally explode as supernovae. Supermassive black holes have masses of millions of solar masses or more and are thought to exist at the centers of many galaxies, including our own Milky Way galaxy.

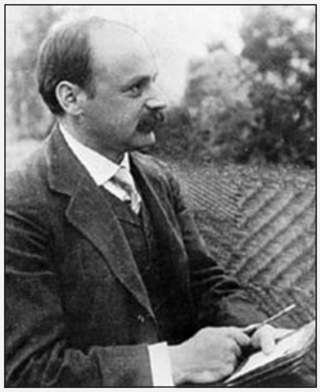
The application of the equations of general relativity to cosmology and the structure of the universe allows for the universe to be flat, negatively curved, or positively curved. Any curvature is related to how the average density of matter and energy in the universe compares to a value called the critical density, which is equivalent to about three hydrogen atoms per cubic meter.

In the simplest form of relativistic cosmology, any possible curvature of the universe is linked to its ultimate fate. If the average density of matter and energy equals the critical density, the universe is flat and it will continue to expand, but slowing down in such a way that it just barely expands forever. In this case the universe is said to be critically open. If the average density of matter and energy is less than the critical density, the universe is negatively curved and it will easily expand forever, slowing

The Schwarzschild Radius

At the start of World War I in 1914, Karl Schwarzschild, a young professor at the University of Göttingen, volunteered for military service. Craving action, he eventually managed to get transferred to Russia, where he heard of Albert Einstein's new general theory of relativity. Schwarzschild wrote two papers on that theory, both published that year. He provided a solution—the first to be found—to the complex partial differential equations fundamental to the theory's mathematical basis. Schwarzschild solved the Einstein equation for the exterior space-time of a spherical nonrotating body. This solution showed that there is an enormous, virtually infinite, redshift when a body of large mass contracts to that certain radius—a size now known as the Schwarzschild radius.

The value of that size is easily calculated by a simple astrophysical formula Schwarzschild derived, relating the radius to the universal gravitational constant, the star's mass, and the speed of light: $R = 2GM/c^2$. Surprisingly, he showed that the general theory of relativity gave basically the same results as Isaac Newton's more common theory of gravitation, but for different reasons. When the mass of the object is measured in units of the Sun's mass, the Schwarzschild radius is neatly given by three times the ratio of the mass to the Sun's mass, the answer expressed in kilometers: $R = 3 \times M/M(\text{Sun})$. If the Sun were contracted to a radius of 3 kilometers, it would be of the right size to be labeled a "black hole." A body becomes a black hole when it shrinks to a radius of less than the critical radius; at that point, nothing, including light, will have enough energy ever to escape from the body—hence the name "black hole," since no light escapes and anything falling in remains. Earth would have to contract to a radius of approximately one centimeter to become a black hole.



While in Russia, Schwarzschild contracted pemphigus, an incurable metabolic disease of the skin. He was an invalid at home in 1916 when he died. He was forty-two years old. For his service in the war effort, he was awarded an Iron Cross. In 1960, he was honored by the Berlin Academy, which named him the greatest German astronomer of the preceding century.

just a little. In this case, it is said to be open. If the average density of matter and energy is greater than the critical density, the universe is positively curved. It will expand to some maximum size and then contract on itself in what some call the big crunch. In this case, it is said to be closed.

Various measures (mostly very indirect) indicate that the universe is extremely close to being flat, but the amount of matter and energy observed is substantially less than the critical density. This observation has led to the notion that most of the mass in the universe is in the form of "dark matter," which has not yet been detected. Moreover, it seems as if only a small part of the unobserved dark matter can be in the form of conventional matter; most of it must be something exotic and as yet unknown.

A completely unexpected discovery in the mid-1990's was that the expansion of the universe seems to be accelerating. This necessitates using a more general form of relativistic cosmology that includes an extra term involving a parameter called the cosmological constant (or some other modification similar to it). To account for the accelerating expansion of the universe, some unknown energy is required. Called "dark energy," it may represent 70 percent or more of the total matter and energy in the universe.

General relativity with its distortions of space-time has thus become an important tool for understanding the origin, structure, and future of the universe and its contents. Although its basic structure has remained unaltered since 1915, it continues to find applications in such diverse areas as the precession of orbits, the bending of light, gravitational redshifts, black holes, and cosmology. It is truly amazing that an abstract theory concerning the warping of space-time has turned out to be so powerful and useful.

George R. Plitnik and Richard R. Erickson

FURTHER READING

Davies, P. C. W. *Space and Time in the Modern Universe*. New York: Cambridge University Press, 1981. Written in an authoritative and lucid style, this book explores the changing ideas of space and time and their applications in astronomical and cosmological scenarios.

Gribbin, John. *Spacewarps*. New York: Delacorte Press, 1984. Describes the universal implications of Albert Einstein's general theory of relativity in physics and astronomy.

_______. *Timewarps*. New York: Delacorte Press, 1980. A clear and imaginative examination of some of the dramatic questions raised by the new concepts of time.

Misner, Charles W., Kip S. Thorne, and John A. Wheeler. *Gravitation*. San Francisco: W. H. Freeman, 1973. This comprehensive, massive textbook is still the fundamental reference on gravity, space-time, and general relativity. It begins with an exposition of the space-time concept of special relativity in a form that facilitates its natural extension to general relativity. Suitable for the diligent layperson with some background in physics, though many later parts are written at more advanced mathematical levels. Its mathematical rigor and sophistication is partly offset by its almost colloquial, folksy style of presentation.

Rucker, Rudolf. *Geometry, Relativity, and the Fourth Dimension*. New York: Dover, 1977. A highly readable and amusing exposition of four-dimensional space-time and the structure of the universe.

Wheeler, John Archibald. *A Journey into Gravity and Spacetime*. New York: Scientific American Library, 1999. Explores different phenomena to explain Einstein's theories of space-time and gravity. A good introductory book for the general reader.

Will, Clifford. *Was Einstein Right? Putting General Relativity to the Test*. 2d ed. New York: Basic Books, 1993. The renaissance of relativity is described with splendid clarity by one of the professional participants. Observations and theories that test the experimental basis for general relativity are presented without mathematics.

See also: Big Bang; Cosmic Rays; Cosmology; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; General Relativity; Interstellar Clouds and the Interstellar Medium; Milky Way; Space-Time: Mathematical Models; Universe: Evolution; Universe: Expansion; Universe: Structure.

Space-Time: Mathematical Models

Category: The Cosmological Context

In both the special and general theories of relativity, neither space by itself nor time by itself is independent of the state of the observer. Only a certain mathematical union of them, called space-time, has invariant properties. The geometry of space-time is the basis for relativistic physics, which are seen in our solar system. A full description of the advancement of the perihelion of Mercury, for example, requires the use of relativity. Also, the way the mass of the Sun can bend light coming from other stars and galaxies is described by relativity theory.

OVERVIEW

Space-time is a four-dimensional coordinate system or reference frame in which one mathematically describes the spatial location and temporal coordinate of an event. Such a frame of reference can be either inertial or non-inertial. An inertial frame of reference often is defined in Newtonian mechanics as a frame that is not accelerated, but then one must ask, "Accelerated relative to what?" A better definition consistent with relativity is that a local inertial frame (or LIF) is a frame in which a body subject to no external force moves at constant speed in a straight line, or alternatively, that it is a frame in which everything is weightless. The word "local" is added because in the space-time of general relativity, it is not possible to have a truly universal inertial frame, if any mass is present.

A wide variety of experiments have repeatedly confirmed that physical phenomena do not fundamentally differ from one inertial frame of reference to any other. The special theory of relativity asserts this as a basic postulate or principle: All the laws of physics are the same in every inertial frame of reference. This means that both the mathematical form of fundamental equations of physics and the values of the physical constants that they contain are the same in all inertial frames. When this principle is applied to the theory of electromagnetism, it requires that observers in all inertial frames of reference agree about the numerical value of the speed of electromagnetic waves in empty space. The universality of this speed—henceforth referred to as the "speed of light" and represented by the letter c—requires that space and time separately cannot be invariant but must change upon transformation from one inertial frame of reference to another.

An event is the name given to something that happens at a particular time and place. The collection of all events (the "whens" and "wheres") in the history of a particle is called its world line. Measurements of the time and place of an event will vary from one inertial frame to another, but there are equations called Lorentz transformations that relate the time and space coordinates of an event in one inertial frame to the time and space coordinates in another inertial frame, based on the relative velocity of the two frames. All the famous phenomena predicted by the special theory of relativity (relativity of simultaneity, length contraction, time dilatation) can be derived from these transformations.

Although the space and time coordinates of an event will differ from one inertial frame to another, there is a quantity called the space-time separation (or space-time interval) between the events that is invariant, meaning that it has the same value in all inertial frames. Let A and B be two events with time and space coordinates (t_A, x_A, y_A, z_A) and (t_B, x_B, y_B, z_B) in one inertial frame, and (t_A', x_A', y_A', z_A') and (t_B', x_B', y_B', z_B') in another inertial frame. The space-time separation s between the two events is defined as

$$s^2 = -c^2 (t_B - t_A)^2 + (x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2$$

as calculated in one frame, and

$$s^2 = -c^2 (t_B' - t_A')^2 + (x_B' - x_A')^2 + (y_B' - y_A')^2 + (z_B' - z_A')^2$$

as calculated in the other frame. It can be shown using the Lorentz transformations that these two expressions for s are equivalent, demonstrating that the space-time separation is invariant between inertial frames.

The interval or separation between two events in space-time is somewhat analogous to the distance r between two points in space:

$$r^2 = (x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2$$

However, there is a fundamental difference between the geometry of space-time and the geometry of space by itself. Notice that the square of the difference in the time coordinates appears in the space-time formula with the opposite sign from the squares of the differences in the spatial coordinates.

APPLICATIONS

Because the squares of the time and space coordinate differences have opposite signs in the formula for space-time interval or separation, the square of the interval or separation between two distinct space-time events can be positive, zero, or negative, depending on how the squared difference of the time coordinates compares to the sum of the squared differences of the space coordinates.

If the squared interval or separation is positive, the squared difference of the space coordinates dominates over the squared difference of the time coordinates, and the separation is called space-like. This means that there exists some inertial frame of reference in which the time coordinates of the two events are equal, so in this frame the two events are simultaneous and they differ only in spatial location. In particular, neither event can be the cause or effect of the other, since all physical influences require time to propagate. In other inertial frames, the events will have different time coordinates, and it is possible for either of the two events to have the larger time coordinate, meaning it occurred later. Consequently, "later" and "earlier" have no universal meaning for a pair of events with a space-like separation, since in some frames one event occurred later while in other frames the other event occurred later.

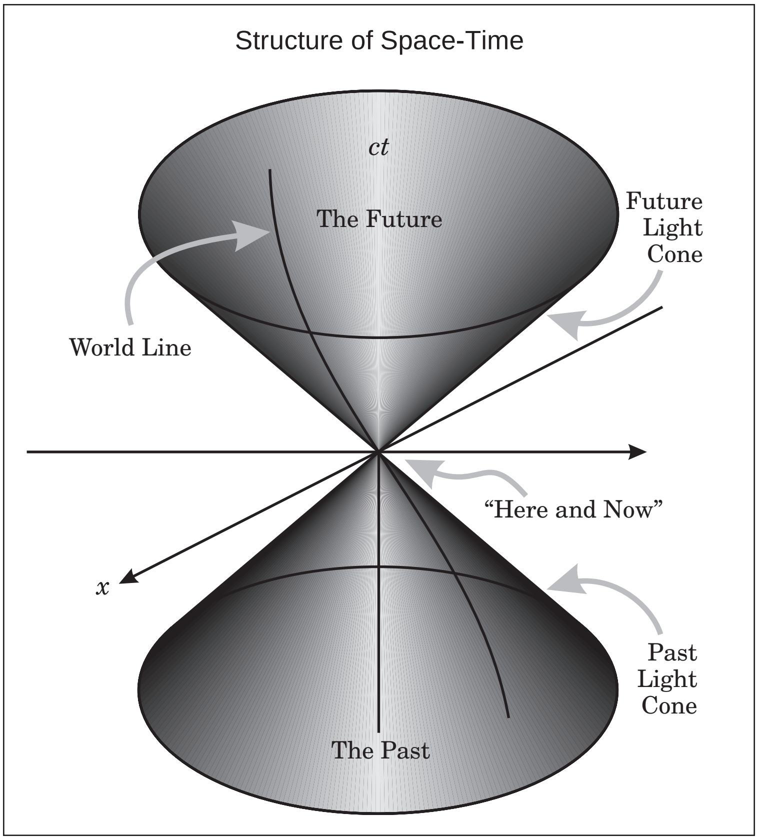
If the squared interval or separation is negative, the squared difference of the time coordinates dominates over the squared difference of the space coordinates, and the separation is called time-like. This means that there is some inertial frame in which the spatial coordinates of the two events are equal, so in this frame the two events occurred at the same location and differ only in time. When a pair of events have a time-like separation, observers in all inertial frames agree on which event occurred first and which occurred second.

If the squared interval or separation is zero, the squared difference of the space coordinates equals the squared difference of the time coordinates (multiplied by the speed of light squared), and the separation is called null or light-like. This means that, in every inertial frame, the pair of events may be connected by the world line of a ray of light moving from one to the other. Such a ray of light could be the agent by which the earlier event causes the later event. Watching a pair of events that have light-like separation, observers in all inertial frames agree on which event occurred first and which occurred second.

Since a particle is always at its own location, intervals between events on the world line of a particle with mass must be time-like. On the other hand, intervals between events on the world line of a photon (a "particle" of light) must be light-like or null.

The sign of the squared intervals between events can be used to divide space-time into regions of different character. Suppose that event A is at the coordinate origin ("here and now") of space-time ($x_A = 0, y_A = 0, z_A = 0, t_A = 0$). As an aid to visualization, the z coordinate may be suppressed; then it is possible to draw a diagram illustrating these regions. The surface mapped out by all null intervals is a double cone. The upper cone (positive t) represents all events in the future that can be reached by a light ray emitted here and now, while the lower cone (negative t) represents all events in the past that could have sent a light ray to arrive here and now. All events that occur within the cone have squared space-time separations from event A that are negative or time-like.

Consequently the world line of any particle with mass that passes through ("coincides with") the event chosen as the "here" and "now" origin of the figure is confined to the interior of the light cone. All events outside the cone have



of relativity which is usually called the twin paradox. Effects of the special theory of relativity are analyzed from inertial frames of reference but may include accelerated objects, such as the twin who travels out and back, thus aging less than the twin remaining at rest in one inertial frame throughout the other's trip.

Analysis of motion from the point of view of observers in accelerated frames of reference is also possible, but it uses the mathematical concepts of differential geometry. As seen from accelerated frames of reference, the structure of space-time is not globally covariant but only locally covariant. This means that the light cones at various events may be tilted in relation to each other. Albert Einstein's general theory of relativity attributes such distortion of the geometry of space-time to gravity. This theory can be summarized in two intimately linked statements: (1) Matter warps space-time, and (2) warped space-time tells matter and light how to move.

squared space-time separations from event A that are positive or space-like. Consequently they can neither influence nor be influenced by event A, for to do so the influence would have to travel between event A and any event outside the cone at a speed greater than the speed of light.

The time axis of any space-time coordinate plot indicates the passage of time in the frame of reference with those coordinates. In an inertial frame, the time axis will be a straight line. The world line, curved or straight, of any particle can be considered a time axis for that particle. Intervals along its world line define its "proper time," which elapses on a clock carried by the particle.

The straight line between two points in Euclidean space has the shortest length of any curve joining them. The geometry of space-time is such that the straight line between two events with a time-like separation has the longest proper time of any world line joining them. This is the basis of a straightforward prediction

CONTEXT

The root of the concept of space-time was the discovery by Hendrik Antoon Lorentz, published in 1898, of the rules of transformation of the coordinates of an event from one inertial frame of reference to any other inertial frame of reference. His derivation was carried out to find a transformation that does not change the form of the fundamental laws of electrodynamics, known as Maxwell's equations. However, Lorentz did not claim that the transformations he found which kept electromagnetism invariant had the broad applicability they are now understood to have. It remained for Einstein to formulate a comprehensive view, published in 1905 in his special theory of relativity, of space and time as measured in inertial frames moving relative to each other, and their dependence on the state of motion of an observer.

Even as he was establishing the foundations of what is now called the special theory of relativity, Einstein was aware of the incompatibility of these ideas with Newtonian gravitational theory. His early work on extending the principle of relativity beyond inertial frames of reference was hampered by mathematical complexities. Hermann Minkowski, a former math professor of Einstein at Zurich Polytechnic University, in an address presented in 1908, introduced the concept of unified space-time based on the new ideas expressed in Einstein's 1905 description of his special theory of relativity. Minkowski realized that Einstein's assertion of the constancy of the speed of light for all observers implied that space and time were fundamentally linked. He also introduced the powerful techniques of geometry, which provided both the mathematical formalism for dealing with noninertial frames as well as intuitive insights into the physical implications. By exploiting ideas first introduced to understand the differential geometry of curved surfaces, unified space-time became the natural way to understand all of physics. Thus, the mathematics of non-Euclidean geometry found application in Einstein's general theory of relativity, which was published in 1916. The general theory is a comprehensive synthesis of the relations among space, time, matter, and motion from the point of view of any frame of reference whatsoever.

In contemporary physics, space-time is accepted as the arena in which all things exist and move. The assertion that the laws of physics must be independent of arbitrary choice of a particular frame of reference in space-time is a powerful working tool of the theoretical physicist. This requirement puts limits on possible new hypotheses and the equations they imply in almost all areas of physics. The one branch of physics that has remained at odds with relativity theory is quantum mechanics. It remains to be seen whether quantum behavior can be unified with the space-time of relativity.

John J. Dykla

FURTHER READING

Ferington, Esther. *The Cosmos*. Alexandria, Va.: Time-Life Books, 1988. This profusely illustrated large-format book is a brief but surprisingly comprehensive introduction to modern cosmology. Includes illuminating presentations on the concept of space-time in both the special and the general theories of relativity.

Hawking, Stephen, and Roger Penrose. *The Nature of Space and Time*. Princeton, N.J.: Princeton University Press, 2000. A collection of lectures given by Hawking and Penrose, who pick up where Niels Bohr and Albert Einstein left off. Technical; for readers with strong physics and mathematics backgrounds.

Hawking, Stephen, Kip Thorne, Igor Novikov, Timothy Ferris, and Alan Lightman. *The Future of Spacetime*. New York: W. W. Norton, 2003. A collection of essays including a basic introduction to the concepts of relativity. Topics discussed include wormholes, gravity waves, and time travel. For the advanced reader.

Minkowski, Hermann. "Space and Time." In *The Principle of Relativity*. New York: Dover, 1952. This early presentation of the space-time concept to a technical audience uses for the most part elementary mathematics to explore some implications of the relativistic unity of space and time.

Misner, Charles W., Kip S. Thorne, and John A. Wheeler. *Gravitation*. San Francisco: W. H. Freeman, 1973. This comprehensive, massive textbook is still the fundamental reference on gravity and general relativity. It begins with an exposition of the space-time concept of special relativity in a form that facilitates its natural extension to general relativity. Suitable for the diligent layperson with some background in physics, though many later parts are written at more advanced mathematical levels. Its mathematical rigor and sophistication are partly offset by its almost colloquial, folksy style of presentation.

Rabinowitz, Avi. *Warped Spacetime, the Einstein Equations, and the Expanding Universe*. New York: Springer, 2009. This college text derives Einstein's equations using calculus-based physics. It also discusses space-time, black holes, worm holes, and cosmology. Designed for undergraduate and graduate science students.

Schwinger, Julian. *Einstein's Legacy: The Unity of Space and Time*. New York: Scientific American Books, 1986. A lively and well-illustrated account of the special and general theories of relativity, this volume includes a careful elementary presentation on the concept of space-time with helpful examples and applications.

Siegfried, Tom. "It's Likely That Times Are Changing." *Science News* 74, no. 6 (September 13, 2008): 26-28. Traces the changing views of time, from Einstein and Minkowski to the present.

Taylor, Edwin F., and John A. Wheeler. *Space-Time Physics*. San Francisco: W. H. Freeman, 1966. This brief text is a thorough treatment of the subject at an elementary mathematical level. Careful reading and working through its examples develops intuition for thinking relativistically. The few problems that require calculus for their solution are clearly identified.

See also: Big Bang; Cosmic Rays; Cosmology; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; General Relativity; Interstellar Clouds and the Interstellar Medium; Milky Way; Space-Time: Distortion by Gravity; Universe: Evolution; Universe: Expansion; Universe: Structure.

Stellar Evolution

Category: The Stellar Context

Stars go through a series of changes that are referred to as stellar evolution or stellar life cycles, in analogy to the life stages of living organisms. Stars are "born" in interstellar clouds of gas and dust called nebulae. They generate energy by various mechanisms for most of their "lives." They "die" when they finally run out of ways to produce any more energy.

OVERVIEW

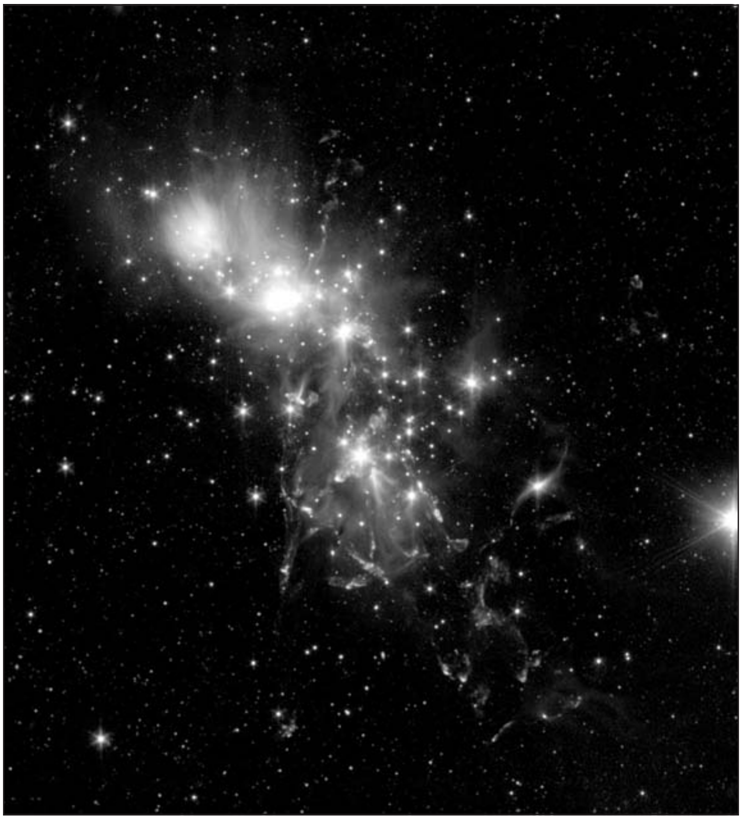
For most of the twentieth century, one of the primary goals of astronomers was to determine how stars are born, how they live, and how they die. By the end of the century, a fairly complete picture finally had emerged. It is convenient to express most stellar parameters relative to the Sun: The Sun's mass is 1.99×10^{30} kilograms (330,000 times the mass of Earth), its radius is 696,000 kilometers (109 times larger than Earth's), its surface temperature averages about 5,800 kelvins, and its luminosity (rate of energy output) is 3.9×10^{26} watts, or 3.9×10^{26} joules per second.

Stars are born in nebulae, large clouds of gas and dust in interstellar space. The clouds typically are several tens of light-years in diameter and contain enough matter to form hundreds to thousands of stars. The gas is mostly hydrogen (about 71 percent by mass), with some helium (about 27 percent by mass) and small amounts of heavier elements. The dust particles are small, solid grains of carbon, silicate minerals, iron compounds, and ices, probably about 0.001 millimeter in size, on average. The nebulae in which stars form are cold, with temperatures on the order on 10 kelvins. At such low temperatures, atoms can bond together to form molecules, so these nebulae are referred to as molecular clouds. Also as a result of the low temperatures, the gas pressure is very weak and can barely keep the cloud from contracting from its self-gravity. In denser parts of the cloud, self-gravity may overcome the weak pressure and start that part of the cloud collapsing. The trigger for this increased density may be an encounter with a spiral arm density wave as the nebula orbits the center of its galaxy, strong stellar winds from a nearby star that already formed, or shock waves from a nearby supernova explosion.

A molecular cloud usually breaks up into separate clumps that collapse on their own. As each clump collapses, gravitational energy is rapidly converted into thermal energy, heating the gas and causing it to begin feebly emitting radio waves. When a sufficiently dense core has formed at the center of the clump, the contraction slows and the object is called a protostar. Usually many neighboring protostars form at about the same time. A protostar continues to contract gravitationally, growing hot enough to begin to shine in the infrared part of the electromagnetic spectrum. What little visible light the protostar emits is blocked by the shroud of dust that surrounds it. Eventually the dust shroud dissipates, some of it joining the growing equatorial disk around the protostar (in which planets may ultimately form) and some of it blown away as the protostar becomes hotter and brighter. As the protostar continues to contract and get hotter, collisions break molecules apart into individual atoms and electrons are stripped off the atoms, ionizing the gas.

When the temperature at the center of the contracting protostar reaches a few million kelvins, hydrogen nuclei start to fuse together to form helium nuclei. The energy released by this nuclear fusion reaction stops the contraction, and the protostar becomes a main sequence star. It takes higher-mass protostars less time to reach the main sequence stage than lower-mass protostars, because their greater mass means stronger gravity and more rapid contraction. The protostar stage lasts on the order of a hundred thousand years for a star with 10 solar masses, a few tens of millions of years for a star like the Sun, with 1 solar mass, and from several hundred million to a billion years for a star with 0.1 solar mass.

In hydrogen fusion, four hydrogen nuclei fuse together to form one helium nucleus. The mass of the four hydrogen nuclei is slightly greater than the mass of the one helium nucleus produced, and the excess mass is converted into energy according to Albert Einstein's formula $E = mc^2$, which states energy E and mass m are equivalent and are related by a physical constant c (the speed of light squared). To overcome the electrical repulsion that the positively charged hydrogen nuclei (bare protons) have for each other requires high temperatures (so the nuclei are moving quickly) and high densities (so the nuclei are close together). These are the conditions in the cores of main sequence stars.



The nebula NGC 1333 in the Perseus constellation, seen here in this infrared image from the Spitzer Space Telescope, shows the birth of new stars "hatching" from the dust clouds in which they formed. (NASA/JPL-Caltech/R. A. Gutermuth, Harvard-Smithsonian CfA)

A main sequence star is in a state of hydrostatic equilibrium, which means the self-gravity trying to make the star contract is balanced by the pressure trying to make the star expand. The energy released by hydrogen fusion in the core provides the energy that the star radiates into space. The main sequence stage is the longest and most stable part of a star's energy-producing life. A star remains a main sequence star as long as it has hydrogen in its core to fuse to helium. Massive main sequence stars have more fuel, but they consume that fuel much more rapidly, so their main sequence lives are relatively short. For example, a 30-solar-mass star has a main sequence lifetime of about 5 million years. The Sun, with 1 solar mass, has a main sequence lifetime of about 10 billion years. (Since the Sun is currently about 4.5 billion years old, it has progressed about halfway through its main sequence life.) Low-mass main sequence stars have less fuel, but they consume it much more slowly, so their main sequence lives are much longer, from about 30 billion years for a star with 0.5 solar mass to as much as a trillion years for a star with 0.1 solar mass. These main sequence lifetimes are longer than the age of the universe, so every low-mass main sequence star that ever formed is still a low-mass main sequence star.

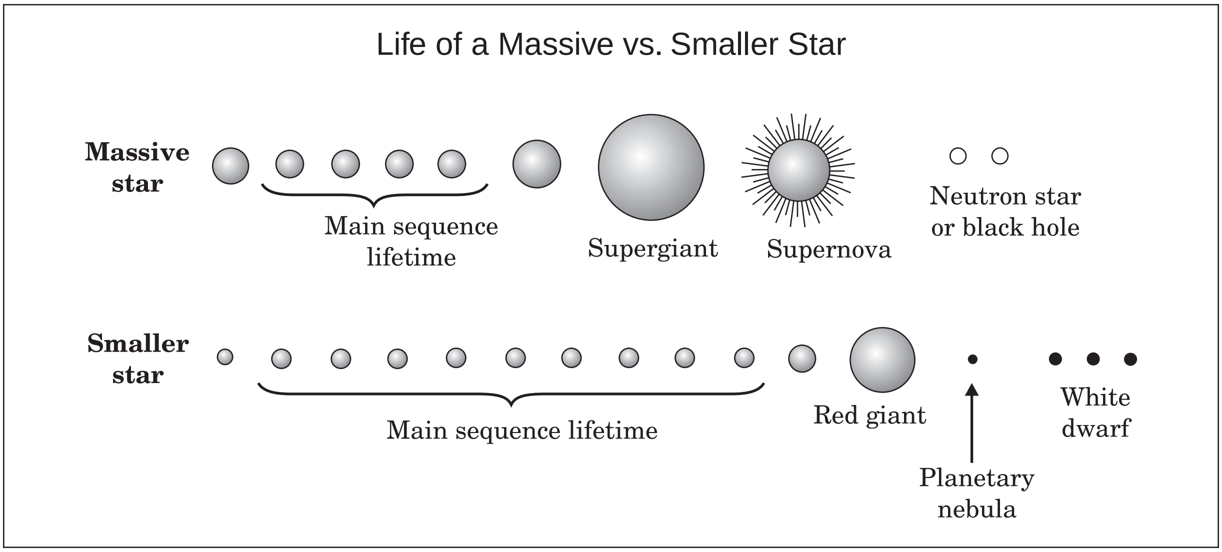
When all the hydrogen in the core has been fused into helium, the core contracts and heats up. Hydrogen fusion is transferred to a still hydrogen-rich shell surrounding the contracting helium-rich core. This causes the outer layers of the star to expand and cool, and the expanding surface of the star becomes red in color. The star ceases to be a main sequence star and becomes a red giant or red supergiant. Stars similar to the Sun, with about 1 or 2 solar masses, become red giants, ten to one hundred times bigger than the present Sun. Massive stars become red supergiants, one hundred to one thousand times bigger than the present Sun.

When the temperature of the contracting helium-rich core reaches about 100 million kelvins, helium fusion is ignited. Three helium nuclei fuse to form one carbon nucleus. Add another helium nucleus, and an oxygen nucleus forms. The total mass of the three or four helium nuclei is a bit greater than the mass of the carbon or oxygen nucleus, and again this mass excess gets converted into energy. This fusion reaction supplies the star with energy for much less time than hydrogen fusion did, when it was a main sequence star.

When the helium in the core is exhausted, the core once again contracts and heats up. A star like the Sun is not massive enough for its core to shrink enough to get hot enough to start any more nuclear fusion reactions. Thermal pulsations and a strong stellar wind blow off its outer layers in one or more shells of gas. The expanding shell of gas surrounding what is left of the star is called a planetary nebula. (Planetary nebulae have nothing at all to do with planets. The name originated in the 1800's because, with the telescopes in use then, they looked round, like planets, and fuzzy, like nebulae.) The central star of a planetary nebula is the exposed former core of the star. It contracts as much as it can and becomes a white dwarf star composed of carbon and maybe oxygen. White dwarf stars have about the mass of the Sun packed into a sphere about the size of the Earth. This makes them very dense, averaging about 1 metric ton per cubic centimeter. White dwarf stars can no longer generate energy; they cannot contract to release gravitational energy, and thus they cannot get any hotter to be able to tap other nuclear fusion reactions. They shine only because they are very hot. As they shine, they radiate their energy away and cool off, gradually becoming black dwarfs (cold, dark, dead stellar "embers").

The core of a massive star (more than about eight solar masses), because of the star's stronger self-gravity, can shrink enough to get hot enough to go through a series of nuclear fusion reactions, one after another, producing heavier atomic nuclei. This stops with the production of iron nuclei, since iron is the heaviest nucleus that can form through fusion reactions that release energy. To form heavier nuclei through fusion requires the input of energy. The iron core collapses, and the outer layers collapse on top of it and rebound, sending shock waves through the star. This tears the star apart in a Type II supernova explosion. In a few minutes, it releases more energy than it produced by nuclear fusion reactions during its entire preceding life. A Type II supernova becomes about a billion times more luminous than the Sun before it gradually fades away. Some of the prodigious energy released in the explosion goes into fusion reactions forming elements heavier than iron. Much of the star's matter is violently ejected into interstellar space at speeds ranging from thousands to a few tens of thousands of kilometers per second, thereby enriching the interstellar material in elements heavier than helium.

If the mass of the stellar remnant that remains after the Type II supernova explosion is less than approximately 2 to 3 solar masses, it becomes a neutron star. A neutron star has a radius of only 15 kilometers and a density of a billion metric tons per cubic centimeter. If the remaining mass of the supernova remnant is greater than about 3 solar masses, however, no



known force can stop its collapse to a black hole. The gravitational field of a black hole is so great that nothing, not even light, can escape from it. Consequently, a black hole cannot be directly "seen" in any part of the electromagnetic spectrum. A black hole can be detected only by its effects, primarily through its gravity, on nearby objects.

KNOWLEDGE GAINED

The Earth and all life on it exist only because of the life and death of massive stars. Our current understanding of the big bang (the primordial "explosion" that created the universe about 13 to 14 billion years ago) is that it produced only hydrogen and helium with very small traces of lithium and beryllium. The first stars that formed in the earliest days of our galaxy consisted only of hydrogen and helium. Gas planets like Jupiter might have formed around those first stars, but not rocky/metallic planets like Earth, nor could carbon-based life have developed. The atoms of all the elements heavier than helium that make up our bodies and the Earth we live on were produced by nuclear fusion reactions in massive stars before and during their deaths as Type II supernovae. These heavy elements were spewed out into interstellar space by the supernova explosions, where they enriched the interstellar clouds of gas from which new stars formed. By the time the Sun and solar system began to form in one of these nebulae about 4.5 billion years ago, about 2 percent of its mass consisted of elements heavier than helium. This provided the material for rocky/metallic planets to form in the inner solar system and for carbon-based life-forms to develop on one of them, our Earth. The production and dispersal of heavy elements by massive stars, with the resulting enrichment of nebulae, continues to the present time, so that today about 5 or 6 percent of interstellar matter consists of elements heavier than helium.

Our understanding of the evolution of stars like the Sun reveals the future of our solar system, including Earth. About 5 billion years from now, the Sun will run out of hydrogen in its core. It will expand until its radius is about one hundred times larger than it is now, swallowing the planet Mercury and possibly Venus. At that point it will be a red giant, about a thousand times more luminous than it is now. Temperatures on Earth will reach between 1,000 and 2,000 kelvins; our atmosphere will escape into space, our oceans will boil away, surface rocks will at least partly melt, and life will not survive here. Eventually, a strong solar wind and thermal pulses will blow the Sun's outer envelope back into space, producing a planetary nebula. What is left of the Sun will become a white dwarf, with perhaps one-half to two-thirds of its present mass. Initially, this white dwarf Sun will have a surface temperature of about 100,000 kelvins, but since it will be unable to generate more energy, it will cool, rapidly at first, then ever more slowly, until it becomes a cold, dark, dead black dwarf.

CONTEXT

The twentieth century witnessed the development of our modern understanding of stellar evolution. New and improved instrumentation, including detectors in space above Earth's atmosphere, made it possible to observe the solar system and deep space with sharper resolution and in all parts of the electromagnetic spectrum. Advances in theoretical physics and the widespread use of computers led to the construction of more and better models of the stellar interiors and the study of how they change with time.

In the late 1890's and early 1900's, George Ellery Hale oversaw the design and construction of several large telescopes—the 40-inch refractor at Yerkes Observatory in Wisconsin, and the 60-inch and 100-inch reflectors on Mount Wilson in California. For the next half century, they were the largest telescopes in the world, and their use led to many observational discoveries in all branches of astronomy.

The early 1900's also saw the development of the Hertzsprung-Russell (or H-R) diagram, a graph used for plotting stellar luminosity versus surface temperature. It turns out that the location of a star in this diagram indicates many things about the star; besides its luminosity and surface temperature, these include its radius, in some cases its approximate mass, and its evolutionary stage. Consequently, the H-R diagram has proved to be a powerful tool for tracing the life cycles of stars of various masses.

In 1926, Sir Arthur Stanley Eddington published *The Internal Constitution of the Stars*, a book that established much of the formalism still used today to construct models of stellar interiors. Its most serious deficiencies, of which Ediington was well aware, involved the sources of stellar opacities and energy generation. In the 1930's, Subrahmanyan Chandrasekhar used relativity theory to work out the structure of white dwarf stars, which, since their discovery in the middle to late nineteenth century, had at best been a stellar enigma and which many astronomers considered to be a physical impossibility.

Continuing through the 1930's and into the 1940's, the problem of stellar energy generation began to be addressed by advances in nuclear physics, which led to the idea that stars generate energy through nuclear fusion reactions; a few specific fusion reactions were suggested for main sequence stars. Work in nuclear physics also suggested the existence of neutron stars, though many thought that even if they did exist, they could never be detected. (It was not until 1967 that the discovery of pulsars, identified as rapidly rotating, highly magnetic neutron stars, confirmed the existence of neutron stars.)

The late 1950's saw the publication of a number of seminal articles and books. A long article by Margaret Burbidge, Geoffrey Burbidge, William A. Fowler, and Fred Hoyle titled "Synthesis of the Elements in Stars" (1957) traced the various processes by which heavy elements could be produced by nucleosynthesis in stars. Martin Schwarzschild's book *Structure and Evolution of the Stars* (1958) described in an accessible way the methods for computing models of stellar interiors. About this time, electronic computers become more readily available, and Schwarzschild's book spurred increased use of them to rapidly calculate many models for the interiors of stars of different masses and in different evolutionary stages. His basic methods can still be applied today to personal computers and even programmable calculators.

The year 1957 also saw the dawn of the space age with the launch of Sputnik 1, the first artificial Earth satellite, by the Soviet Union. Visible light is distorted as it passes through Earth's turbulent atmosphere, resulting in images that are somewhat blurred. One way to get around this problem and produce sharper images is to observe from satellites above our atmosphere. The Hubble Space Telescope, placed in orbit at an altitude of 600 kilometers in 1990, has supplied incredible high-resolution images of various stages of stellar evolution. Since observing time on space telescopes is limited and sparingly assigned, techniques have been developed for ground-based telescopes to provide sharper images. In speckle interferometry, many short exposure images (with little blurring due to atmospheric turbulence, since the exposure is so short) are recorded electronically and then combined in a computer to produce a sharper final image. In adaptive optics, sensors monitor the effects of atmospheric turbulence and control small motors that slightly alter the telescope optics to compensate for it and reduce the blur. These techniques have given new life to large ground-based telescopes.

Observations from space are essential, however, for observations in certain nonvisible portions of the electromagnetic spectrum, because Earth's atmosphere is opaque to much of the electromagnetic spectrum, including gamma rays, X rays, and most ultraviolet and infrared radiation. Gamma rays, X rays, and ultraviolet radiation are emitted by hot objects; infrared radiation is emitted by cool objects. It is necessary to observe in all these wavelength regions to study both the cool births and the hot deaths of stars. Placing astronomical instruments on board satellites orbiting outside Earth's atmosphere has allowed astronomers to observe over all parts of the electromagnetic spectrum. Infrared observations are revealing regions of star formation and the embryonic stars developing there. Around many of them it is possible to detect disks in which planets might eventually form. Ultraviolet and X-ray observations have revealed that stars lose much more mass during the giant and supergiant stages of their lives than was previously thought, so computer models of stellar evolution are being revised to account for this discovery.

George E. McCluskey, Jr.

FURTHER READING

Asimov, Isaac. *The Exploding Suns: The Secrets of the Supernovas*. New York: E. P. Dutton, 1985. An excellent book, by the well-known physicist and author of science fiction and popular science books, which describes how massive stars die. Discusses the evolution of stars and explains how some of them reach their ultimate fate—the cataclysm of a supernova explosion. A highly accessible treatment.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Contains a thorough description of the stages of stellar evolution, complete with transparent overlays.

Cohen, Martin. *In Darkness Born: The Study of Star Formation*. New York: Cambridge University Press, 1988. The author, a recognized authority in the field of star formation, has written the first general survey of the subject. The book discusses how stars form from interstellar material and how astronomers can use both ground-based and space-based astronomical instruments to observe this intriguing process. Appropriate for the person with little scientific background. Includes many black-and-white plates.

Cooke, Donald A. *The Life and Death of Stars*. New York: Crown, 1985. Offers general background on the subject and provides an overall picture of the Galaxy. There follow chapters on how stars form, how they shine and change as they burn their fuel supplies, and finally how they die as all energy sources are depleted. For the general reader.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages Through the Universe*. 3d ed. New York: Brooks/Cole, 2006. A well-written, thorough college textbook for introductory astronomy courses. Includes a good basic overview of stellar evolution.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook offering a good description of stellar evolution.

Genet, Russell M., Donald S. Hayes, Douglas S. Hall, and David R. Genet. *Supernova 1987a: Astronomy's Explosive Enigma*. Mesa, Ariz.: Fairborn Press, 1988. The full story of the brightest supernova in almost four hundred years: Supernova 1987a, discovered in March, 1987. This book discusses the evolution of this massive star from birth to its explosive demise in a supernova explosion. The book, which has more than one hundred illustrations, introduces the reader to the discoverers of Supernova 1987a and explains how astronomers used both ground- and space-based observatories to study this event.

Greenstein, George. *Frozen Star*. New York: Charles Scribner's Sons, 1983. Devoted to neutron stars, pulsars, and black holes, this book is clearly written and contains many diagrams. The structure of a neutron star is clearly illustrated, and the discussion of black holes is very helpful. To illustrate these bizarre objects, the author takes his readers on an imaginary trip to a pulsar and discusses what would happen if the Sun became a black hole.

Kippenhahn, Rudolf. *One Hundred Billion Suns: The Birth, Life, and Death of the Stars*. New York: Basic Books, 1985. Kippenhahn, who helped to pioneer the computer calculation of stellar evolution in the 1960's, has written an authoritative review of the subject of the life cycle of stars. The book is written in a delightful style and is one of the best treatments of the subject. Well illustrated.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses. Divided into lots of short sections on specific topics. Contains a thorough discussion of the stages of stellar evolution.

See also: Brown Dwarfs; Gamma-Ray Bursters; Main Sequence Stars; Novae, Bursters, and X-ray Sources; Nuclear Synthesis in Stars; Protostars; Pulsars; Red Dwarf Stars; Red Giant Stars; Supernovae; Thermonuclear Reactions in Stars; White and Black Dwarfs.

Sunspots

Category: The Sun

Sunspots are small areas on the solar photosphere that appear darker than their surroundings because they are not as hot. They occur at sites with strong magnetic fields, which are their cause. Sunspots vary in number over an eleven-year cycle, and their associated magnetic fields switch polarity with each eleven-year cycle, so the full magnetic cycle is twenty-two years.

OVERVIEW

Sunspots appear as darker spots on the brighter photosphere of the Sun. Their temperature is about 1,000 to 1,500 kelvins cooler than the rest of the photosphere (which is about 5,800 kelvins); as a result, sunspots are only about one-third as bright, so they look dark in comparison. The spots have diameters ranging from a few hundred kilometers to occasionally more than 100,000 kilometers. Generally, smaller spots last only a few days, while larger ones linger for several weeks. At any given time, the Sun may display lots of sunspots or it may have none at all. Their numbers cyclically increase and decrease with a period of approximately eleven years. Especially large sunspots could occasionally be seen with the unaided eye, although the Sun should never be observed with the naked eye or directly through devices such as cameras or telescopes; such observation will severely damage the eyes.

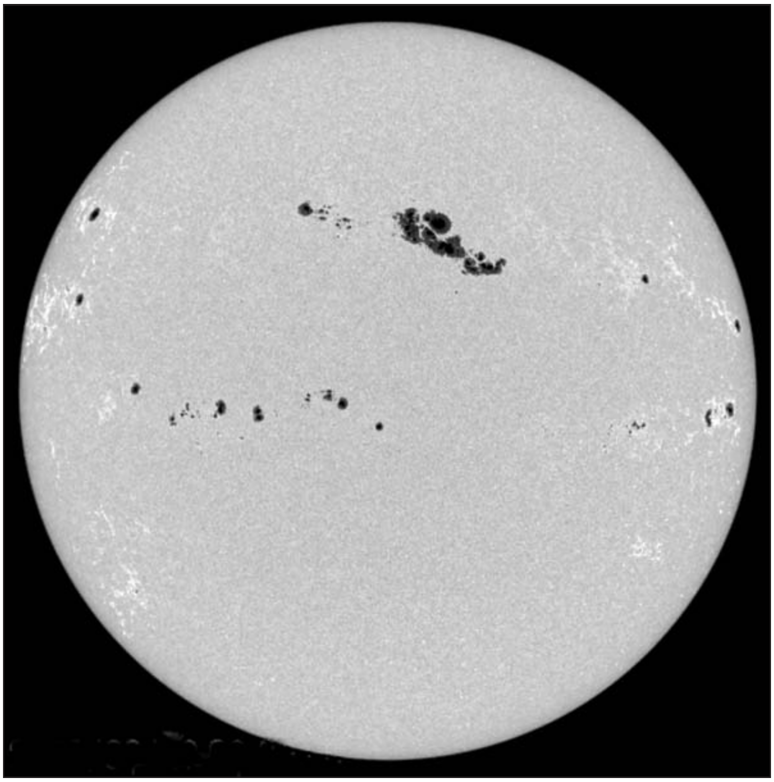
In 1843, Samuel Heinrich Schwabe (1789-1875), a German pharmacist and amateur astronomer, announced after seventeen years of recording sunspots that the number of spots increased and decreased over an eleven-year cycle. At the beginning of each sunspot cycle, spots first appeared at latitudes around 40° north and south of the solar equator, progressing down to latitudes around 5° north and south at the end of that cycle; then the next cycle began.

A few years later, in 1859, the British astronomer Richard Carrington found that the motion of sunspots across the solar disk depends on their latitude. He concluded that the rotation period of the Sun varies with latitude, ranging from about 25 days near the equator to about 28 days at 40° north or south latitude. (This latitude range is the band in which sunspots are commonly seen. We now can measure the Sun's differential rotation in other ways, and have found that it increases to about 33 days at 75° north or south latitude.)

In 1908, George Ellery Hale, the American astronomer responsible for establishing Yerkes, Mount Wilson, and Palomar observatories, found that absorption lines in the spectra of sunspots were split into two or more components. He attributed this to the Zeeman effect, which describes the splitting of spectral lines in the presence of strong magnetic fields. It had been noted that sunspots usually occur in close pairs, one ahead of the other in the direction of solar rotation. Hale discovered that the two spots in a pair have opposite magnetic polarities; that is, the magnetic field lines coming out of one spot return into its neighbor. Moreover, the leading spot in all the spot pairs in one solar hemisphere (north or south of the equator) have the same magnetic polarity, and the polarity of the leading spots is opposite in the other solar hemisphere. In 1924, Hale announced that the magnetic polarities of spot pairs reverse with each eleven-year sunspot cycle, so the complete magnetic cycle is twenty-two years, twice as long as the sunspot cycle.

Each sunspot has a central darker region called the umbra, which is surrounded by an outer not-as-dark region called the penumbra. The temperature is as low as 4,000 to 4,500 kelvins in the umbra and averages 5,000 to 5,500 kelvins in the penumbra, compared to around 5,800 kelvins in the adjoining photosphere; consequently the umbra is only about one-third as bright as the photosphere, so it looks dark in comparison. Within the central umbral region of the sunspot, the magnetic field is on the order of 3,000 gauss (0.30 tesla), strong enough to deflect away hot bubbles of ionized gas rising from the bottom of the Sun's convection zone up toward the photosphere. This inhibits the transfer of heat from below, making the spot cooler and thus darker. Furthermore, the cooler, denser gas in the umbra sinks, drawing in gas from inside and outside the surrounding penumbra.

The diameter of the umbra seldom exceeds 20,000 kilometers (although an exceptionally large one might on rare occasions exceed 100,000 kilometers), limiting its depth to about the same distance and large-scale mass movement to about half its size. Sunspots and their associated fields appear to be anchored at this depth. Vertical flow of matter downward in the umbra is found to be limited to 25 meters per second, while the horizontal flow from the penumbra into the umbra does not exceed 50 meters per second. Sunspots also show irregular patterns of bright points called "umbral dots" or "umbral granulation," in addition to solar filigrees caused by delicate, small-scale movement of magnetic field elements.

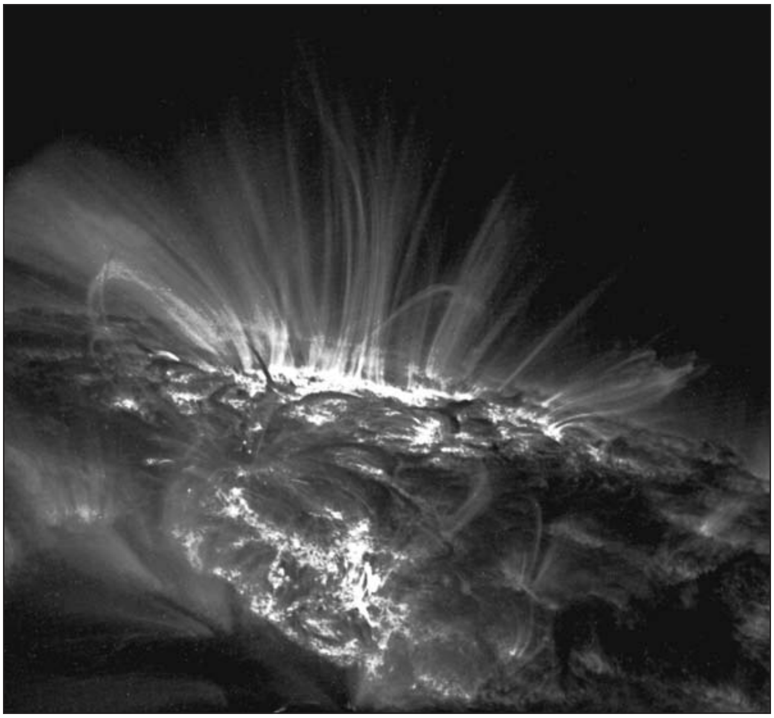


Some very large sunspots can be seen in this 2001 image from the Solar and Heliospheric Observatory. (NASA)

The average magnetic field strength in the penumbra is about one-half the field of the central umbra. The penumbral magnetic field has fine horizontal structures, giving the region a filamentary overlapping white and gray appearance. Penumbral matter flow is nearly horizontal, inward toward the umbra, with progressively decreasing velocities farther out.

"Pores," smaller regions with dark umbrae, also occur, having magnetic fields on the order of 2,000 gauss (0.20 tesla). "Magnetic knots" are compact magnetic structures, having fields on the order of 1,000 gauss (0.1 tesla).

The dynamo theory of solar magnetic fields was developed by Eugene N. Parker, Horace W. Babcock, and others in the 1950's to explain the Sun's magnetism and associated phenomena, such as sunspots, flares, and prominences. The outer zone of the Sun's interior, below its visible surface or photosphere, is in a state of convective motion; the convective cells that form are responsible for the observed boiling and bubbling (called granulation) on the solar surface. Along the granule boundaries, tubular magnetic fields emerge and disperse into the solar atmosphere. Basically, it is theorized that the rising and sinking motions of the ionized gas (plasma) in the convective zone constitute electrical currents that induce magnetic fields through a dynamo process. These magnetic fields get distorted and intensified by the Sun's differential rotation. As described earlier, the solar rotation period is about 25 days at the equator, progressively increasing to 28 days at 40° north and south latitudes, and 33 days at 75° north and south latitudes. The general background dipole magnetic field, produced by convective motion and "frozen" into the ionized gas of the convection zone, gets twisted and wound up around the Sun by this differential rotation, thus increasing the local magnetic field strength.



Sunspots on the solar surface are surrounded by glowing arcs of gas. (NASA)

At the beginning of an activity cycle, the Sun's magnetic field is generally that of a dipole, with magnetic field lines running from near one rotational pole to the other. Differential rotation, after a few rotations, stretches these field lines that are frozen into the ionized gas near the surface so that they wrap around the Sun parallel to the equator. As the magnetic field lines are wound more closely together, the local field strength increases. The convective motion of the plasma further twists the field lines like strands of rope. The increased intensity of the magnetic field lines at higher latitudes finally makes them break through the solar surface as sunspots roughly along the east-west direction. As the field weakens at higher latitudes by the dissipation of energy through sunspots, the field intensifies toward the equator, forcing the locations of new sunspots to migrate toward the equator. The sunspot cycle ends with the last spots fading near the equator and the magnetic field there weakening.

The next solar activity cycle begins with the dipole field reversed from the preceding activity cycle. Thus the solar magnetic cycle has a period double that of the activity cycle. The solar dynamo theory satisfactorily explains the incidence and location of sunspots during each activity cycle, and the reversal of spot polarity in the two solar hemispheres from one activity cycle to the next. However, the complexity of the proposed mechanisms warrants further investigation.

APPLICATIONS

The sunspot cycle seems to be related to long-term temperature fluctuations on Earth. There appear to have been long episodes of solar inactivity, as evidenced by historical records of reduced numbers of sunspots and other related manifestations, such as smaller coronas seen during total solar eclipses and fewer auroras. These generally occurred during times when the Earth's climate was cooler. An accumulating body of information indicates that the Sun may be a slightly variable star, possibly with complex periodicity. Such a prospect for solar luminosity changes would have far-reaching ramifications for life on Earth.

It is believed that the magnetic fields of the major planets, including the Earth, are products of a dynamo mechanism similar to that operating in the Sun. The study of sunspot cycles led to the solar dynamo theory, which in turn has contributed to progress in understanding planetary magnetic fields.

The Sun is the nearest star. Everything learned about the Sun can be extended to most other stars. Spectroscopic studies of magnetic fields in other stars have revealed that many stars possess fields similar to those found on the Sun, with intensities ranging from 2,000 to 20,000 gauss (0.20 to 2.00 tesla). The observed stellar dipolar fields appear to vary, with periods ranging from half a day to decades. Magnetic field activity, presumably similar to that of the Sun, appears to be stronger in younger stars, which rotate more rapidly in comparison with older and more slowly rotating stars. For example, the younger stars in Orion have magnetic field intensities three times that of similar-sized older stars.

Among many stars, "starspots" are suspected to exist. They tend to occur mainly in relatively younger stars. It is possible that sunspot activity and associated magnetic field intensity may have been stronger at an earlier time in our solar system, when the Sun may have been spinning more rapidly.

CONTEXT

Recorded descriptions of sunspots date back to at least the fourth century B.C.E. In 1612, Galileo (1564-1642) was the first to observe them with a telescope. He noticed their positions on the Sun's bright disk changed over a few days, and from this he concluded that the Sun rotates about once a month. These observations, along with many others Galileo made with his telescopes, contradicted the prevailing views held then by scholars throughout Europe that dated back to the ancient Greeks.

Today, much has been learned about the Sun and its physics through both Earth-based and, more important, space-based observatories and instruments. Solar research is concerned with many broad areas of the Sun's structure and observable activity, such as sunspots and the solar magnetic cycle. Solar activity has been determined to be periodic, marked by the occurrence of sunspots and related phenomena, and caused by complex variations in associated magnetic field intensities. It has become possible to study the Sun at virtually all wavelengths of the electromagnetic spectrum (from high-energy gamma rays to low-frequency radio waves) with the aid of telescopes on the ground and spacecraft outside the Earth's atmosphere. The advent of supercomputers has permitted very detailed modeling of the structure and development of sunspots and other solar activity. Advances in plasma physics and magnetohydrodynamics continue to refine our understanding of the solar dynamo mechanism and the related activity cycle, including sunspots.

V. L. Madhyastha

FURTHER READING

Balogh, André, Louis J. Lanzerotti, and S. T. Suess, eds. *The Heliosphere Through the Solar Activity Cycle*. New York: Springer, 2007. Investigates the Sun during the sunspot cycle to analyze correlations between the cycle and the heliosphere. Discusses theory on the subject as well as actual data from spacecraft.

Brody, Judit. *The Enigma of Sunspots: A Story of Discovery and Scientific Revolution*. Edinburgh, Scotland: FlorisSunspots, 2002. Provides an overview of the history of sunspot research, from earliest times to the present.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley,

2008. Very well-written college-level textbook for introductory astronomy courses. Good description of sunspots and the Sun's magnetic field.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Good description of sunspots and the Sun's magnetic field.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook that includes a good description of sunspots and the Sun's magnetic field.

Gibson, Edward G. *The Quiet Sun*. NASA SP-303. Washington, D.C.: Government Printing Office, 1973. Although somewhat out of date, this NASA publication contains many useful photographs, illustrative graphs, and charts. Technical, but written so that a nonscientist will not experience any difficulty with the text.

Giovanelli, Ronald G. *Secrets of the Sun*. New York: Cambridge University Press, 1984. This book, which contains many photographs of sunspots and their progression in the eleven-year cycle, is specially written for scientists and lay readers alike. A rare example of a serious attempt to introduce the reader to a complex set of ideas.

Jordan, Stuart, ed. *The Sun as a Star*. NASA SP-450. Washington, D.C.: Government Printing Office, 1981. Although technical, this volume is a complete source of solar physics. Includes a large number of references at the end of each chapter. A nontechnical reader can skip over the occasional mathematical equations without experiencing a sense of loss in logic or continuity.

Maunder, E. Walter. "A Prolonged Sunspot Minimum." *Knowledge* 17 (1894): 173-176.

_______. "The Prolonged Sunspot Minimum, 1645-1715." *Journal of the British Astronomical Society* 32 (1922): 140ff. These two papers are Maunder's original publications arguing that there was a prolonged minimum in sunspot activity in the late seventeenth century. Together these papers form a seminal work on sunspots.

Newkirk, Gordon, Jr., and Kendrick Frazier. "The Solar Cycle." *Physics Today* 35 (April, 1982): 25-34. Using the solar magnetic-dynamo model, the authors discuss eleven-year and twenty-two-year solar cycles. Considers a possible clock in the Sun, the variation of luminosity correlated with sunspot activity, and photospheric pulsation.

Parker, E. N. "Magnetic Fields in the Cosmos." *Scientific American* 249 (August, 1983): 44-54. The author, an authority on the subject of cosmic magnetic fields, presents in precise, qualitative terms the theory of the solar dynamo, the mechanism that, combined with the differential rotation of the Sun with latitude, affords a natural explanation of the intense magnetic fields associated with sunspots, flares, and other aspects of the sunspot cycle.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. This introductory astronomy textbook offers several sections on sunspots and the Sun's magnetic field.

Wentzel, G. Donat. *The Restless Sun*. Washington, D.C.: Smithsonian Institution Press, 1989. This excellent volume contains photographs, illustrations, and information on all types of solar activities. Geared for beginners and nonspecialists.

See also: Auroras; Coronal Holes and Coronal Mass Ejections; Earth's Magnetic Field at Present; Earth-Sun Relations; Solar Chromosphere; Solar Corona; Solar Evolution; Solar Flares; Solar Geodesy; Solar Infrared Emissions; Solar Interior; Solar Magnetic Field; Solar Photosphere; Solar Radiation; Solar Radio Emissions; Solar Seismology; Solar Structure and Energy; Solar Ultraviolet Emissions; Solar Variability; Solar Wind; Solar X-Ray Emissions.

Supernovae

Category: The Stellar Context

Supernovae are stars that explode violently at the end of their lives. They are spectacular events, emitting light at rates up to ten billion times greater than the Sun. They are responsible for the synthesis of chemical elements heavier than iron and dispersing them through interstellar space. Shock waves from supernovae can act as triggers for new star formation.

OVERVIEW

Novae and supernovae have been observed and recorded since ancient times as "new" stars that appeared suddenly in the night sky (hence the name "nova" which is from the Latin *novus*, "new"). Originally there was no distinction between novae and supernovae, since superficially both display the same behavior: a nova or "new star" suddenly appears in the sky, brightening dramatically in no more than a few hours, and then slowly fades from view over a period of weeks to months. It was not until the 1920's that astronomers began to realize there was a difference; some of the outbursts were much brighter intrinsically than the others, up to a million or more times. The "super" outbursts came to be called supernovae. Today it is known that both novae and supernovae are not new stars but explosive events involving stars in the last stages of their lives. A nova is a relatively minor hydrogen fusion explosion on the surface of a white dwarf star, while a supernova is a much more violent explosion that tears a star apart.

A supernova is one of nature's most spectacular events. The total energy released is on the order of 10^{46} joules. Most of the energy is carried by a flood of neutrinos produced in the explosion, while a much smaller part goes to drive the explosive outburst that destroys the star. The electromagnetic radiation emitted by the blast is only on the order of 10^{43} joules, but this is almost as much as the Sun emits during its entire life. At its peak, a supernova shines as bright as 100 million to 10 billion Suns.

Supernovae are rare events. The last one to occur in our Milky Way galaxy was observed in the year 1604. Thus historical records of supernovae are invaluable resources for modern astronomy, even though those who recorded their observations long ago had no real understanding of what they witnessed. Descriptions of past supernovae can reveal when and where they occurred, how bright they got, and how long they remained visible in the sky. Their frequency (or rather infrequency) provides estimates of the death rates among stars.

Chinese records of celestial events date back to several hundred years B.C.E. They believed events in the sky were omens that forecast events on Earth, and woe to the astronomer who failed to correctly warn of an impending catastrophe. They recorded "broom stars" (comets with tails that "swept" the sky like brooms) and "guest stars" (novae and supernovae that suddenly appeared and eventually faded slowly away, like visiting guests). "Guest stars" were sighted and recorded in what would have been these years (all C.E.) in our current calendar system: 185, 386, 393, 1006, 1054, 1181, 1572, and 1604.

The supernova that appeared in May, 1006, was particularly bright. More than twenty reports of it have been identified from all around the world. Because of its position in the southern sky and thus below the horizon in northern Europe, the best records of it come from Chinese, Japanese, and Middle Eastern sources. It was a brilliant spot of light perhaps one hundred times brighter in appearance than Venus, visible during the day and bright enough to cast shadows at night. A gaseous remnant of the explosion, now very faint, was identified in the 1970's in the constellation of Lupus, the wolf. The rate of expansion of its filaments indicates it is about a thousand years old, consistent with its being a remnant of the supernova of 1006.

The supernova that appeared in July, 1054, was probably brighter than Venus (though not as bright as the supernova in 1006), visible in the daytime sky for nearly a month and in the night sky for two years. Interestingly, this bright object, although surely visible, went unmentioned in European historical records, possibly because its appearance in the sky somehow conflicted with church doctrine in medieval

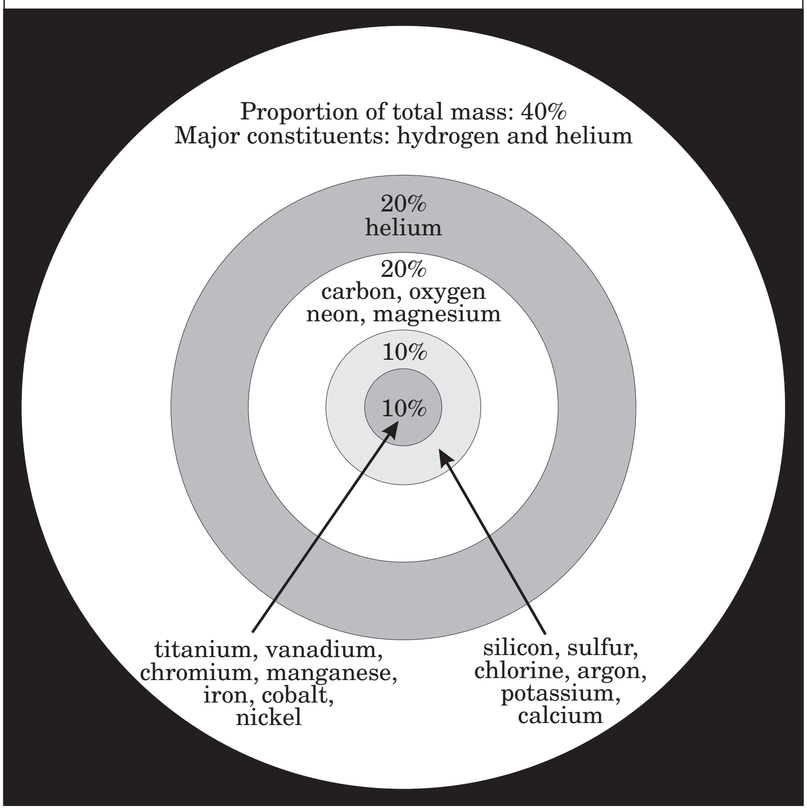
Europe. The event definitely was recorded in China and the Middle East. A rock painting made at approximately that time by Native Americans in New Mexico shows a ten-pointed object next to a crescent and may represent this supernova when it first appeared next to a thin, waning crescent Moon. The descriptions of the position of this supernova match approximately the location of the Crab nebula in the constellation of Taurus the bull. The measured expansion rate of the Crab nebula is consistent with its origin in the supernova of 1054. This supernova left behind, in the center of the Crab nebula, a rapidly spinning neutron star called a pulsar.

On November 11, 1572, a supernova brighter than Venus appeared in the constellation of Cassiopeia, gradually fading from view over the next eighteen months. Most European scholars at that time still adhered to the teachings of Aristotle and Plato that the heavens were perfect and unchanging, and therefore this bright new thing could not be a star but had to be something close to Earth, perhaps in the air overhead. The young Danish astronomer Tycho Brahe made careful measurements of its position for as long as it was visible. Since it showed no motion relative to other stars, he concluded it was far from Earth, like the stars, but had to be a new star no one had seen before. (Brahe's work so impressed King Frederick II of Denmark that he financed the construction of two observatories for Brahe on the Danish island of Ven. It was from these observatories that Brahe measured—without a telescope, since the telescope had not yet been invented—the changing positions of the planets that Johannes Kepler used to derive his three laws of planetary motion.) An expanding shell of gas ejected by the 1572 supernova and plowing into the surrounding interstellar gas has been detected at visible, radio, and X-ray wavelengths. No central stellar object has been found, suggesting that the star was completely destroyed by the explosion.

Kepler, who had been Brahe's assistant for a year prior to his death in 1601, sighted a supernova on October 17, 1604. He described the star as similar to Jupiter in brightness, with the color of a diamond. It remained visible through October, 1605, although the star briefly disappeared behind the Sun from November, 1604, until January, 1605. Sightings of this supernova were also recorded in China and Korea. A remnant from this supernova was first photographed in 1947. This was the last supernova to be observed in the Milky Way.

Fortunately, because supernovae are so bright, hundreds have been observed in galaxies outside our own. Modern measurements of

Anatomy of a Supernova



their properties combined with historical records of past supernovae have yielded our current understanding of supernovae.

Several types of supernova spectra can be identified. Type II supernova spectra have emission lines of hydrogen, while Type I supernova spectra do not. Type I spectra are subdivided into Type Ia, with a strong absorption line of ionized silicon; Type Ib, with a strong absorption line of neutral helium; and Type Ic, with neither of these absorption lines. There are two distinct ways that supernovae explode. Types II, Ib, and Ic are core-collapse supernovae, in which the iron cores of massive supergiant stars collapse, triggering the supernova explosion. In contrast, Type Ia supernovae occur in white dwarf stars in close binary systems that acquire so much matter from their companions that their mass is raised above the Chandrasekhar limit (the maximum mass a white dwarf star can have); the whole white dwarf then collapses and explodes.

To understand how stars explode as supernovae, one must understand the earlier stages of their lives. Stars form from nebulae, clouds of gas and dust in interstellar space, by gravitational contraction. When the center of a contracting protostar reaches a temperature of a few million kelvins, hydrogen nuclei begin fusing into helium nuclei in the star's core, releasing energy. This stops the contraction, and the star becomes a stable main sequence star. It remains a stable main sequence star as long as it has hydrogen in its core to fuse into helium. This is the longest-lasting stage in a star's life. (The Sun is a main sequence star; it has been one for most of the 4.5 billion years of its existence, and it probably has enough hydrogen left in its core to last several billion years more.) When hydrogen is exhausted in its core, the star expands to become a red giant or supergiant and begins to fuse helium into carbon and maybe oxygen in its core.

Stars with up to about 8 solar masses cannot go further in nuclear fusion reactions. When their core helium is exhausted, they puff off their outer layers as expanding shells of gas called planetary nebulae (a purely descriptive term, having nothing to do with planets), and their cores become electron-degenerate white dwarfs (meaning their electrons are packed as tightly as quantum mechanics allows), shining only because they are hot. As they radiate their energy away, they grow cooler and fainter, eventually becoming cold, dark black dwarfs. This usually is the end of life for stars of this size.

However, a white dwarf in a close binary star system can end its life in a much more dramatic way if its binary companion transfers enough matter to the white dwarf. Since white dwarfs are electron-degenerate, there is a maximum mass they can have, the Chandrasekhar limit, which equates with about 1.4 solar masses. If enough matter is transferred from the companion star to the white dwarf so its mass is raised above the Chandrasekhar limit, its self-gravity exceeds the pressure of the degenerate electrons and it collapses in on itself. The rapid compression drives the temperature up to billions of kelvins, igniting a whole series of nuclear fusion reactions that release tremendous amounts of energy. This is a Type Ia supernova. At its peak, it shines as brightly as nearly 10 billion Suns. The star apparently is completely destroyed, its matter flung into interstellar space as a cloud of rapidly expanding gas containing the heavier chemical elements produced in the fusion reactions. The companion may become a runaway star, no longer held in orbit by the mutual gravitational attraction between the two stars. The last two supernovae seen in the Milky Way—in 1572 and 1604—apparently were Type Ia.

A different kind of supernova explosion, called a core-collapse supernova, is in store for stars more massive than about 10 solar masses. They are massive enough to continue to form heavier elements through various nuclear fusion reactions while they are supergiant stars. Each element formed in one fusion reaction becomes the fuel for the next reaction. Eventually the supergiant builds up a small central core of iron surrounded by shells of lighter nuclei, similar to the thin layers in an onion. In each of these shells, nuclear fusion reactions form heavier nuclei that become part of the next shell inward, except that iron is the heaviest nucleus that can form in fusion reactions that release energy. To form heavier nuclei than iron requires the input of energy. The iron core shrinks, trying to tap another nuclear fuel, but none is available. The temperature climbs as high as a few hundred billion kelvins, causing the iron nuclei to begin to break up. When the central density reaches about 400 billion times the density of water, free electrons are forced into the nuclei, where they combine with protons to form neutrons and neutrinos. The loss of free electrons reduces the pressure so much that the core collapses at speeds as fast as one-fourth to one-third the speed of light, shrinking from a diameter about as large as the Earth to less than 20 kilometers in less than one second. The outer layers collapse inward on top of the core, where some of the matter rebounds and collides with the still infalling gas above. This collision heats the gas to billions of kelvins, igniting fusion reactions that release energy. Even more energy is added by the outward flow of neutrinos. Although neutrinos do not ordinarily interact very much with other forms of matter, the density near the core is so great that the neutrinos transfer some of their energy to the collapsing outer layers, reversing their fall and blasting them outward at speeds greater than 10,000 kilometers per second. This core-collapse supernova has a spectrum of Type II, Ib, or Ic. (The differences in the spectra depend on how much of their outer layers the stars lost as a result of strong stellar winds while they were supergiants.) The huge amounts of energy now available allow fusion reactions to occur that form the chemical elements heavier than iron. The outer layers become a rapidly expanding supernova remnant that disperses into interstellar space the heavy elements formed both in the supergiant before it exploded and in the supernova explosion itself.

If the neutron-rich core at the center of this explosion is less than about 3 solar masses, it becomes a neutron star, a neutron-degenerate sphere about 20 kilometers in diameter with an average density of about a billion metric tons per cubic centimeter or about 10^{15} times the density of water. (Neutron degeneracy means that the neutrons are packed as tightly as quantum mechanics allows.) Neutron stars typically rotate very rapidly, up to several hundred times per second, a consequence of conserving their angular momentum as they collapsed. Their magnetic fields become concentrated, with two very strong magnetic poles that do not necessarily lie at the two rotational poles. Electromagnetic radiation beamed outward from the two magnetic poles can appear to blink on and off as seen from Earth in a "lighthouse beacon" effect as the star rotates, if the rotation and magnetic axes are aligned the right way. Such blinking neutron stars are called pulsars, and their blinking can be detected throughout the electromagnetic spectrum, from radio waves (where they were first discovered) to visible light to X rays. There is a pulsar spinning 30 times per second at the center of the Crab nebula, produced by the Type II supernova in 1054. The fastest-spinning pulsar known rotates more than 600 times per second.

However, if the core at the center of the explosion exceeds about 3 solar masses (called the Oppenheimer-Volkoff limit), neutron degeneracy will not stop its ultimate collapse to a black hole. "Black hole" is the name given by the physicist John Wheeler in 1969 to matter with such strong gravity that nothing, not even light, can escape from it. Although black holes cannot be observed directly in any part of the electromagnetic spectrum, precisely because no form of electromagnetic radiation can escape from them, they can be detected by the gravitational effects they have on matter nearby. Several likely stellar-mass black holes, the kind that can be formed from supernovae, have been identified, and it is strongly inferred that a black hole lies at the center of the Milky Way.

The final mass of the core, whether it is below or above the Oppenheimer-Volkoff limit, is what determines whether a neutron star or black hole is left by the supernova explosion. Computer models of the late stages of massive stars suggest that the final core mass is determined by the mass the star had when it was born. Stars with initial masses less than about 20 to 40 solar masses probably produce neutron stars, while those with initial masses greater than 20 to 40 solar masses probably produce black holes.

APPLICATIONS

Because supernovae emit so much light, they can be observed at very large distances, up to billions of light-years away in very distant galaxies. Core-collapse supernovae (Types II, Ib, and Ic) come from stars with a wide range of large masses and consequently somewhat different interior structures and histories. The explosions are not all identical and thus there is a range of peak luminosities; they do not all become equally bright. On the other hand, almost all Type Ia supernovae reach nearly the same peak luminosity, nearly 10 billion solar luminosities, because they have essentially identical progenitors: white dwarfs with masses that just exceed the Chandrasekhar limit of 1.4 solar masses. (Occasionally a Type Ia supernova may noticeably exceed the normal maximum brightness. This is thought to be due to a collision between the white dwarf and its binary companion.) The vast majority of Type Ia supernovae therefore make very reliable standard candles; their assumed "standard" real brightness, combined with their measured apparent brightness, yields their distance. If a Type Ia supernova is observed in a galaxy, then the distance to the galaxy can be determined. Type Ia supernovae were used to find the distances of galaxies in two independent projects that show the expansion of the universe is accelerating.

Supernova SN 1987A was first seen before dawn on February 24, 1987. It was bright enough to see with the unaided eye, since it was relatively close, only 160,000 light-years away in the Large Magellanic Cloud (LMC), a small galaxy that forms a companion to our own, much larger Milky Way galaxy. SN 1987A provided the first opportunity to study a relatively nearby supernova with a whole array of modern instruments. It was also possible to identify its progenitor in earlier images of the LMC, so astronomers know what kind of star produced it: a blue supergiant with about 20 solar masses. Hence, SN 1987A was a core-collapse, or Type II, supernova. Based on theory and computer models, astronomers formerly thought that core-collapse supernovae occurred in red supergiants; SN 1987 led astronomers to conclude that any massive supergiant could explode as a core-collapse supernova.

Neutrinos from the explosion were detected almost a day before the supernova was seen at visible wavelengths; the neutrinos were able to escape directly from the collapsing core, while the light was absorbed and reradiated by matter between the collapsing core and the star's surface, emerging almost a day later with the supernova shock wave. The time difference from the detection of the neutrinos to the first observed light from the supernova provided information on the size of the star just prior to the explosion, confirming that it was a blue supergiant rather than a red supergiant. Comparing the observed neutrino flux with the observed visible light, astronomers noted that the neutrinos carried more than one thousand times the energy of the visible light. Overall, less than 0.1 percent of the energy of the explosion was emitted in the form of visible light, about 1 percent of the energy was used to propel most of the star's matter outward, and the rest of the energy was carried away by neutrinos.

SN 1987A brightened rapidly at first, becoming a thousand times brighter in less than a day. Then it continued to brighten slowly, eventually reaching about 100 million solar luminosities after about a hundred days. In appearance it looked about as bright as some of the stars of the Little Dipper. Its peak luminosity was not as bright as would be expected for a Type II supernova produced by a 20-solar-mass supergiant. However, a blue supergiant is more compact and not as extended as a red supergiant of the same mass, so more energy would be needed to break up a blue supergiant, leaving less energy to appear as visible light.

For about the first forty days, the energy being radiated came from the explosion itself; afterward it came from the radioactive decay of unstable isotopes created in the explosion. The radioactive decays produced gamma rays that were absorbed by the expanding gas, heating it, and the energy was reemitted at visible wavelengths. Some of the gamma rays were detected directly by instruments in Earth orbit at the energies emitted by radioactive decays of nickel 56 and cobalt 56, isotopes expected in the aftermath of a supernova, thus confirming some theories about supernova nucleosynthesis.

Several months after the initial outburst of SN 1987A, spectroscopic observations with an infrared telescope flown at an altitude of 12,000 meters on a jet transport identified many more expected elements, including iron, nickel, cobalt, oxygen, neon, sodium, potassium, silicon, and magnesium. The intensity of the infrared spectral lines proved that the quantities of these elements were larger than what could have been present initially in the star at birth.

A few years later, a bright ring of gas about one light-year in diameter was observed around the star, and it remained bright for several years. This was not a ring of material ejected by the supernova. Rather it was matter blown off the star by stellar winds about 30,000 years earlier, when it was a red supergiant. The ring lit up when high-speed gas ejected by the supernova plowed into it, compressing and heating the matter in the ring. Meanwhile, an expanding "blob" was observed at the center of the ring, thought to be slower-speed gas ejected by the explosion.

CONTEXT

Supernovae are significant for a number of reasons. Perhaps the most important is that we owe our existence to massive stars that exploded as supernovae long ago. The big bang that created the universe about 13 to 14 billion years ago produced significant amounts only of hydrogen and helium. All the chemical elements heavier than helium—the carbon in our cells, the nitrogen and oxygen in the air we breathe, the silicon and iron in the Earth, the gold and silver in our jewelry, and everything else heavier than helium—were produced by nuclear fusion reactions in massive supergiants and the supernovae that destroyed them. These elements were dispersed into interstellar space by the supernova blasts, where they enriched the clouds of gas from which new generations of stars and accompanying planets formed. The first stars to form after the big bang would have contained only hydrogen and helium. Any planets that developed around them would have been gaseous planets composed of hydrogen and helium, maybe similar to Jupiter and Saturn, but not rocky/metallic planets like Earth. By the time our Sun and solar system formed about 4.5 billion years ago, enough massive stars had already run through their life cycles and exploded as supernovae that the nebula in which our Sun and solar system developed had become enriched in elements heavier than helium, so that they accounted for about 2 percent of its mass. This provided the raw materials for rocky/metallic planets like Earth and for life as we know it.

Second, supernova shock waves propagating through large interstellar molecular clouds may provide the trigger for new bursts of star formation, compressing the gas and starting the gravitational contraction that leads to new stars and solar systems. (Supernova shock waves are only one of several possible ways to initiate the gravitational contraction that forms new stars.)

Supernovae may contribute to the evolution of life on Earth in two ways. Supernovae probably are a major source of high-energy cosmic-ray particles. Trapped by the magnetic field of our Milky Way galaxy, they continually bombard Earth. Although most never reach the Earth's surface, enough do to possibly cause genetic mutations in the life-forms on Earth, resulting in the diversity of life here. Although most genetic mutations are not advantageous and some are outright deadly, life could not evolve and adapt to changing environmental conditions without a steady supply of mutations. If it were not for mutations, then all organisms today would be mere replicas of the initial simple forms of life.

However, if a supernova were to explode too close to us, the large burst of electromagnetic radiation and high-energy particles would pose a serious danger to life. A supernova within 50 light-years of our solar system probably would wipe out all life on Earth, and one within 100 light-years could result in major extinctions. Although the mass-extinction event 65 million years ago that wiped out the dinosaurs and many other forms of life was probably the result of an asteroid impact, other extinctions in the history of Earth might be due to nearby supernovae. A relatively minor extinction of some marine species about 2 million years ago might have been related to a supernova explosion about 120 light-years away. Of course, an extinction of only some species opens up new ecological niches to those organisms that do survive, which in turn can lead to rapid evolution of new species as the survivors rush to fill the environmental holes.

Supernovae are very rare. The last one seen in our Milky Way galaxy was in 1604 C.E. According to estimates, up to five supernovae may occur in the Milky Way Galaxy each century, most escaping detection because they are obscured by interstellar dust in the galactic plane. In other galaxies, the rate varies from several each century in the largest spirals to one every few centuries in the smallest galaxies.

Since supernovae are so rare, it is hard to predict when or where the next one is likely to occur or exactly what the progenitor star will look like. The star will be either a massive supergiant in its last stages before collapse or a white dwarf in a close binary system. Fortunately, the suspected possible candidates are sufficiently far away that they should not pose a threat to life. One of the closest is Canopus (Alpha Carinae), the second brightest star in appearance in the night sky, located about 200 to 300 light-years away. Two other possible candidates are located in the constellation of Orion the hunter: the red supergiant Betelgeuse (Alpha Orionis), about 430 light-years away, and the blue supergiant Rigel (Beta Orionis), about 770 light-years away. Betelgeuse, in particular, seems unstable and pulsates. Another possible candidate is the supergiant star in the Eta Carinae nebula. With a mass of more than 100 solar masses, it is very luminous and is expelling mass at a prodigious rate. It has had a number of outbursts; the one observed in 1843 is the largest single loss of mass that any star is known to have survived.

Michael L. Broyles

FURTHER READING

Asimov, Isaac. *The Exploding Suns*. New York: Dutton, 1985. This famous author of major science books and science fiction covers many aspects of supernovae, from the observations of earlier civilizations to modern theories. An intriguing chapter outlines the formation of elements in a supernova and describes how life may have been affected by such random catastrophes.

Cooke, Donald A. *The Life and Death of Stars*. New York: Crown, 1985. Spectacular photographs and diagrams superbly illustrate all chapters. The Crab nebula is among several case studies highlighted. The bright star in the Eta Carinae nebula is mentioned as a possible future supernova, along with several other stars.

Gribbin, John, with Mary Gribbin. *Stardust: Supernovae and Life, the Cosmic Connection*. New Haven, Conn.: Yale University Press, 2000. The authors explain Carl Sagan's famous assertion, "We are made of star stuff," from dust particles in nebulae through supernova explosions. Also covers the nucleosynthesis theory and the role supernovae play in it. For the general reader.

Gussinov, Oktay, Efe Yazgan, and Askin Ankay, eds. *Neutron Stars, Supernovae, and Supernovae Remnants*. New York: Nova Science, 2007. Describes the core-collapse theory for supernovae, remnants of supernovae that can be found in nearby galaxies, and the connection between neutron stars, supernovae, and remnants.

Marschall, Laurence A. *The Supernova Story*. New York: Plenum Press, 1988. The search for supernovae in the night sky is presented from the early records to those discovered in modern times. A documentary chapter on Supernova 1987A discovery is included, with emphasis on the observational techniques used.

Mezzacappa, Anthony, and George M. Fuller, eds. *Open Issues in Core Collapse Supernova Theory*. Hackensack, N.J.: World Scientific, 2005. Proceedings from the Institute for Nuclear Theory conference on supernovae and the core-collapse theory. Includes papers on computer simulation, quantum chromodynamics, magnetic fields during core collapse, and radiation diffusion. For scientists, graduate students, and advanced undergraduates.

Mobberley, Martin. *Supernovae and How to Observe Them*. New York: Springer, 2007. A how-to book for those intersted in observing supernovae using modest telescopes and charge-coupled devices to record the images.

Murdin, Paul, and Lesley Murdin. *Supernova*. New York: Cambridge University Press, 1985. An excellent collection of photographs and diagrams is included, with discussions of the ancient sightings. Tables are compiled for the magnitudes of Brahe's and Kepler's supernovae. A unique energy-flow diagram for a supernova illustrates the nuclear processes of the evolving star.

North, Gerald. *Observing Variable Stars, Novae, and Supernovae*. New York: Cambridge University Press, 2004. Includes a CD-ROM by Nick James. A guide for amateur astronomers on how to observe supernovae and variable stars. Includes information on telescopes, detectors, and observational techniques. For the general reader.

Wheeler, J. Craig. *Cosmic Catastrophes: Supernovae, Gamma-Ray Bursts, and Adventures in Hyperspace*. New York: Cambridge University Press, 2000. Discusses the birth and death of stars, as well as gravity, wormholes, and supernovae. For the general reader.

See also: Brown Dwarfs; Cosmology; Gamma-Ray Bursters; General Relativity; Hertzsprung-Russell Diagram; Interstellar Clouds and the Interstellar Medium; Lunar History; Main Sequence Stars; Neutrino Astronomy; Novae, Bursters, and X-Ray Sources; Nuclear Synthesis in Stars; Planetary Formation; Protostars; Pulsars; Red Dwarf Stars; Red Giant Stars; Stellar Evolution; Thermonuclear Reactions in Stars; Ultraviolet Astronomy; Universe: Expansion; White and Black Dwarfs.

T

Telescopes: Ground-Based

Category: Scientific Methods

In the age of the Hubble Space Telescope and other space-based observatories, ground-based telescopes are often perceived as antiquated. However, ground-based telescopes are still playing key roles in the search for extrasolar planets and other fields of cosmological research, and larger telescopes are always under construction or development.

OVERVIEW

Ground-based telescopes have a long history of advancing various areas of astronomy research. Starting in the 1960's, large ground-based telescopes (which, for the purposes of this discussion, are those telescopes situated in advantageous locations on Earth's surface and having apertures larger than 250 centimeters) were used to explore and further study objects in our solar system. For example, in the 1950's, the Martian atmosphere was believed to be thin, and, based on the intensity and polarization of Martian reflected light, Martian surface pressure was determined to be only 5 to 10 percent of that on Earth. Designers of preliminary Martian landers used this value when deciding whether a descent to the Martian surface should be effected by balloon, glider, or downward-pointing rocket engine. Also, the presence of this much atmosphere suggested that Martian life might exist, as there would be enough air to breathe (if free oxygen were present) and enough protection from ultraviolet light and extreme temperature changes.

However, a University of California scientist, Hyron Spinrad, using ground-based telescopes, found evidence that the correct value might be very different. When he was able to obtain good high-dispersion spectra using the Lick Observatory's 3-meter telescope in the early 1960's, Spinrad concluded that the true value for the Martian surface pressure must be only about 5 millibars. This is only 0.5 percent of Earth's atmospheric pressure—a very small value. New designs had to be considered for Martian spacecraft, and at that time it seemed less likely that anything remotely resembling terrestrial life could exist under such harsh conditions.

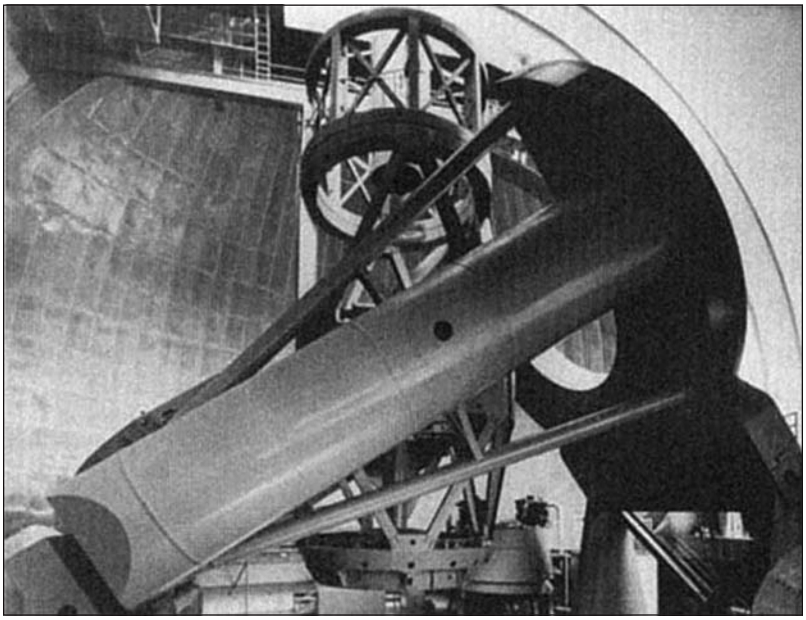
The oldest observatory in the United States is Mount Wilson, which overlooks Pasadena, California. George Hale founded the observatory in 1904 and used funding from the Carnegie Institute of Washington, D.C., to create a solar research laboratory. The Snow Solar Telescope was moved to Mount Wilson in 1904 from Yerkes Observatory in Wisconsin. Long-term studies of the Sun began in 1905. Hale ordered the mirror for his 100-inch telescope in 1906. Two years later, the 60-foot solar telescope became operational. Hale used it when he detected the Sun's magnetic field. In May, 1912, a solar telescope was mounted atop a 60-foot-tall tower, and work began on the foundation for Hale's 100-inch mirror telescope. The latter telescope achieved first light on November 2, 1917, to much fanfare. Several discoveries of great importance in the history of astronomy were made using the telescopes at Mount Wilson. These include detecting the first Cepheid variable star; measuring the distance to M31 (the Andromeda nebula, now called the Andromeda galaxy), thereby proving that the Milky Way is but one of many galaxies in the greater universe; determining extragalactic distances and providing evidence that the universe is expanding; detecting remnants of the 1604 supernova; and discovering four additional satellites of Jupiter.

Mount Wilson remains one of the top observatories conducting astronomical research. In 1995, it became home to Georgia State's CHARA stellar interferometer array. The 150-foot-tall tower telescope is operated by the Astronomy and Astrophysics Division of the University of California at Los Angeles. It is used to study solar magnetic activity and look for long-term changes that may provide an understanding of the Sun's variability and its implications for the solar radiation received by the Earth and its biosphere. The solar telescope is part of a worldwide network monitoring the Sun's surface (its photosphere). The 60-inch telescope, which became operational in 1908, is also the largest in the world available for approved use by the general public.

One of the oldest large optical telescopes in the United States is the Palomar Observatory's 5-meter Hale reflector. Completed in 1948, for thirty years it was the world's largest. Located in the mountains northeast of San Diego, California, Palomar was used primarily for deep space cosmology and stellar studies. With the advent of the space program and the accompanying renaissance of planetary astronomy, the Palomar telescope came to be used occasionally for solar-system work. It was used, for example, to analyze the Martian atmospheric gases and, in the 1950's, to examine the Martian surface for signs of chlorophyll to see if greenish surface regions might indicate plant life; they did not. With infrared detectors, the telescope was used to make some of the first temperature maps of the observed disk of Jupiter and to sample the temperatures of the surfaces of the other major planets.

The Palomar telescope has played an important role in space research with its exploration of objects that were first identified by the Infrared Astronomical Satellite (IRAS) as being anomalous infrared sources. Many of these objects are obscured stars and interstellar clouds within the Galaxy. These are not easily detected and studied at optical wavelengths, but Palomar observers have used infrared detectors, working at short enough wavelengths that are transmitted through Earth's atmosphere. There is still a considerable amount of research to be done before the place of infrared-emitting galaxies in the general scheme of the extragalactic universe is understood thoroughly.

The largest telescope at the Lick Observatory—located at Mount Hamilton, east of San Jose, California—is the Shane 3-meter telescope, completed in 1959. It has been used by University of California astronomers to make numerous discoveries related to space research. An exotic example of such discoveries has to do with the planet Mercury. In 1961, there were reports from the Soviet Union that Mercury appeared to have an atmosphere; according to one Soviet astronomer, it was composed of hydrogen, but another report claimed that it was carbon dioxide. Space scientists were concerned about these reports. Because Mercury's surface was known to have high temperatures, astronomers had deduced that the planet should not be able to retain any appreciable atmosphere, because gases would be so hot that they would escape into space in a relatively short time. In 1962, to explore the question further, the Shane telescope was used to obtain high-dispersion spectra of Mercury. Because Mercury is always quite close to the Sun (never more than 28° from it in the sky), the telescope was used during the daytime for these observations. The results, af-



The 200-inch telescope at Mount Palomar, California. (NASA) ter analysis, were clear. The better data provided no sign whatsoever of an atmosphere. Years later, in 1979, Mariner 10 confirmed this conclusion by returning data that showed that, because of its high temperature, Mercury does not have an appreciable atmosphere. It only retains temporarily a minute and tenuous envelope of hydrogen gas captured from the solar wind.

In 1979, the National Aeronautics and Space Administration (NASA) built a 3-meter infrared telescope to be used especially for space-related research. It was put on Mauna Kea in Hawaii to take advantage of the mountain's high altitude (4,200 meters) and lack of water vapor. It has been employed successfully in a large number of infrared projects, including studies of asteroids, comets, planetary satellites, the galactic nucleus, and infrared-bright and active galaxies.

The 2.7-meter telescope of the McDonald Observatory, near Austin, Texas, was built in 1968 with NASA sponsorship and with the intent that it be used largely for space-related research. A particularly interesting example of this has been its lunar-ranging measurements. A powerful laser at its focal point is used to beam visible light to the lunar surface, and 2.5 seconds later the telescope detects the reflected light. The light's travel time can be measured so accurately that the distance to the Moon could almost instantaneously be determined to an accuracy of about 2.5 centimeters.

In the mid-1980's some astronomers felt that ground-based astronomy had reached its limit. However, advances in computers and technology would prove them wrong. The first telescope using a new technique known as active optics was the New Technology Telescope (NTT) at the European Southern Observatory (ESO); it was installed in 1988. NTT has a 3.58-meter mirror that is only 24 centimeters thick and is therefore flexible. Traditional telescopes hold their shape by the thickness of the mirror. This thickness increases the mirror's weight and limits its size. The shape of NTT's mirror is maintained by a series of supports known as actuators, which are computer-controlled. Success with NTT design led to the construction of ESO's Very Large Telescope (VLT), which is an array of four 8-meter mirrors.

Two of the most famous telescopes using active optics are located at the Keck Observatory in Hawaii near the summit of the dormant Mauna Kea volcano. These telescopes (Keck I and Keck II, completed in 1993 and 1996) each have mirrors measuring 10 meters in diameter. The two can also be linked together to form an even larger interferometer. Keck I and II are used for investigations of brown dwarfs, globular clusters, black holes, Jupiter's atmosphere, and distant young galaxies, for example.

Active optics helps minimize the limiting effects of the telescope itself, increasing light collection and image quality. Adaptive optics, on the other hand, reduces or eliminates the effects of the Earth's atmosphere. Telescopes using adaptive optics (such as Keck) can achieve image resolutions of 30 to 60 milli-arc seconds at infrared wavelengths. Without adaptive optics, they would be able to produce resolutions of only one arc second. One method of adaptive optics uses a guide or reference star. The telescopes are equipped with a wavefront sensor that measures the atmospheric distortion of the light. Information is sent to a computer, which alters the telescope's deformable mirror to correct the image. When astronomers study very distant objects, the target is often too faint for this process. Instead, astronomers use a brighter reference star that is located near the target. Light from the guide star is then used to determine how to adjust the telescope to negate effects of atmospheric distortion of the target object. A second method for adaptive optics is using a laser beam in place of a reference star. The laser guide star (LGS) is directed toward the upper atmosphere and often is pulsed. Reflected light is then detected as it travels back down through the Earth's atmosphere, where it is used as an "artificial" reference star. LGS has advantages for scientists, because a reference star of sufficient brightness is not always found in all parts of the sky.

Technological advances of both active and adaptive optics have led to an ever-increasing number of larger ground-based telescopes. The largest optical telescope in the United States is the Large Binocular Telescope (LBT) located near Safford, Arizona. The telescopes are a joint project with the Italian and German astronomical communities, the University of Arizona, Arizona State University, Northern Arizona University, Ohio State University, Research Corporation of Tucson, and the University of Notre Dame. The LBT consists of two 8.4-meter mirrors with a common mount. Therefore the two telescopes together have the equivalent power of an 11.8-meter telescope. The first telescope was built in Italy and was shipped to Arizona in 2002. In March, 2008, LBT produced its first binocular images, making it the most powerful optical telescope in the world. The first images produced were false-color pictures of NGC 2770, a spiral galaxy 102 million light-years away. One image shows hot young stars concentrated in the spiral arms; the second shows older stars that are more evenly dispersed. The third image is a composite photograph that shows the full range of stars from cool to hot. An instrument is being developed to work with the telescopes at near-infrared wavelengths that would produce images ten times sharper than ones taken by the Hubble Space Telescope. A high-resolution spectrograph is also planned, which would help astronomers study the magnetic field of the Sun and other stars by looking at Zeeman splitting of spectral lines.

Several other extremely large telescopes (ELTs) are being designed or are currently under construction. The 25-meter Giant Magellan Telescope (GMT) is a joint endeavor of the Carnegie Institute of Washington, Harvard University, the Massachusetts Institue of Technology, the University of Arizona, the University of Michigan, the Smithsonian Institution, Australian National University, the University of

These two Keck telescopes at the top of Mauna Kea, Hawaii, can be linked together to form the Keck Interferometer, the world's most powerful ground-based optical telescope. (NASA/JPL)

Texas at Austin, and Texas A&M University. Located in Las Campanas, Chile, GMT will have six mirrors surrounding a central one. Each mirror will measure 8.4 meters, giving GMT a diameter of 25.2 meters. In August, 2008, grinding and polishing of the primary mirror began. When GMT is completed in 2017, scientists will use it to study the origins and evolutions of planetary systems, the formation of black holes, and dark matter.

The Thirty Meter Telescope (TMT) will have 492 hexagonal mirrors, each measuring about 1.44 meters in diameter. The 30-meter telescope will use six laser guide stars to study distant star systems and galaxies. Sites in Chile and Hawaii are being considered, and TMT could be constructed and operational by 2013.

The Large Aperture Mirror Array (LAMA) will combine a series of eighteen 10-meter liquid mirrors. The mirrors of the LAMA telescope are coated with a very thin layer of liquid mercury. The layer is initially 2-3 millimeters thick. The mirrors spin constantly, and after cleaning are started spinning manually. After the mirror is closed, a layer of oxide forms on top, which seals the surface. This allows the extra mercury to be skimmed off and can achieve a liquid mercury layer about 1 millimeter thick. Liquid mirrors are favored over glass ones, because on average they cost 95 percent less. Another benefit of liquid mirrors is the fact that the primary mirror does not tilt. This means that the support structure does not need to be as massive, thus also reducing cost. Together the mirrors that form the array will have the power of a 42-meter telescope. Sites are being considered in New Mexico and Chile. Scientists plan to use LAMA to study distant galaxies, stars, and extrasolar planets.

The most ambitious telescope in development is the European Southern Observatory's 100-meter Overwhelmingly Large Telescope (OWL). The primary mirror will be composed of 3,042 segments, each 1.6 meters in diameter. OWL's spherical primary mirror technology was originally used in the 1990's on the Hobby Eberly telescope. The ESO estimates OWL's cost to be 1.2 billion euros. Locations in Chile, Argentina, the Canary Islands, and near the South Pole are being considered. When fully operational, OWL will be able to observe stars with magnitudes as low as 38 (the faintest stars that can be seen with the unaided eye have a magnitude of 6). OWL will be used to study stellar and planetary system formation, extrasolar planets, and the mysterious dark matter.

Radio telescopes form another category of ground-based telescopes. These telescopes, used to receive and analyze extraterrestrial radiation at radio wavelengths, have been of great importance to the development of modern astronomy and space science. Most large radio telescopes have been one of three distinct types: large single dishes, interferometric arrays of dishes, and millimeter telescopes.

Centimeter-wavelength radio waves can best be detected by means of a very large parabolic reflecting surface (a "dish") that can move in position to follow the celestial object as it moves across the sky. The largest of these actually does not move, however, but instead has a moving receiver suspended above the dish. The 300-meter radio telescope of the Arecibo Ionospheric Radio Observatory was made by smoothing out a depression in the hills of Puerto Rico and lining it with a parabolic metal surface suspended a few feet above the ground. Three pylons rise from the rim of the valley. These hold cables that span the depression and support the small laboratory, dangling above the center of the dish, that contains the receiver. Motors pull the receiver building from west to east to compensate for the diurnal motion of the image of the celestial object being observed. Only a limited range in position in the sky is available to the Arecibo telescope, but its immense collecting area and high resolution have made it an important instrument for many years.

A different type of single-dish antenna is the 100-meter Effelsberg radio telescope near Bonn, West Germany. In this case, the parabolic metal dish is fully movable; it can turn to any place in the sky and can follow celestial sources at the diurnal rate. Although it is smaller than the Arecibo telescope, its flexibility has made it useful for many projects that would have been impossible with a large, fixed dish.

Similar, though smaller, single dishes that have been important are a 76-meter telescope at

Jodrell Bank, England, a 64-meter telescope at Parkes, Australia, and a 43-meter one at Green Bank, West Virginia. A 90-meter transit-type radio telescope at Green Bank collapsed suddenly in 1988; it was replaced by the Robert C. Byrd Green Bank Telescope (GBT), which is the world's largest fully steerable single-dish radio telescope. The 100-meter telescope became operational in 2000.

High resolution at radio wavelengths is hard to achieve. Resolution is proportional to wavelength; to achieve the same resolution afforded by an optical telescope, a radio telescope working at a long wavelength such as the 21-centimeter neutral hydrogen line must be many kilometers in diameter. This size would be impractical for a single-dish design, so radio astronomers have constructed giant arrays of radio dishes, connected electronically so that the signals received are blended and analyzed as if from an immense single telescope. These instruments are called "interferometric arrays" and range in size from 1 kilometer or so for the pioneer instrument in Cambridge, England, to intercontinental arrays.

The largest interferometers that are located in physical proximity are the Very Large Array (VLA) near Socorro, New Mexico, and the Westerbork Array in the Netherlands. The former consists of twenty-seven 25-meter dishes spread out in a Y-shaped pattern on the dry lake bed of the St. Augustine plain; the latter consists of a linear array of telescopes. All radio interferometers of this type have at least some of their antennae on wheels and tracks so that the spacing can be adjusted according to the needs of different observing projects.

The Australia Telescope array was designed with some of the properties of both the VLA and Westerbork; a continental array of telescopes planned to span North America would permit extremely high resolution. Ad hoc interferometers, made up of existing single dishes that are coordinated to simulate an array, have utilized even larger baselines (for example, from Australia to Canada), achieving radio images that are more detailed than even optical telescopes can produce, using normal detectors. Spaceborne radio telescopes can achieve even wider separations, as large as the solar system.

Millimeter-wavelength radio telescopes tend to have characteristics midway between those of standard radio telescopes and those of optical telescopes. Most are housed in domes of some sort and are fully steerable. For example, the 9-meter telescope on Kitt Peak in Arizona has a dome-shaped housing with a large slit that can be opened for observing. Other large millimeter telescopes are located in Hawaii (on Mauna Kea), in Japan, and in Massachusetts. They are especially powerful for detecting and analyzing emissions from cool molecular gas in star-forming regions.

The world's largest millimeter telescope is located on the Sierra Negra volcano, 350 kilometers southeast of Mexico City. The Large Millimeter Telescope (LMT) has a 50-meter dish. LMT is a joint effort between the National Institute of Astrophysics, Optics, and Electronics (of Mexico) and the University of Massachusetts. Construction was completed in 2006, after which followed a two-year testing period before initial research began. Scientists plan to use the large telescope to study extrasolar planetary atmospheres, compositions of comets, and the origins of the universe.

First conceived in 1995, the Atacama Large Millimeter/Submillimeter Array (ALMA) is a joint effort of organizations from North America, Europe, Japan, and Chile. ALMA is located in the Atacama Desert of the Andes Mountains in Chile. When completed in 2012, the array will include up to eighty radio antennas operating at 0.3 to 9.6 millimeters, at a cost of $1.3 billion. The dishes range in size from 7 to 12 meters in diameter. ALMA should produce resolutions about ten times better than the Hubble Space Telescope. Radio astronomers will use ALMA to study galactic nuclei, quasars, distant stellar compositions, and galactic, stellar and planetary formation.

The largest radio telescope presently in development is the Square Kilometer Array (SKA). When finished, this international facility will be fifty times more sensitive than any existing radio telescope. SKA will consist of thousands of radio dishes measuring 10 to 15 meters in diameter. Sites in South Africa and Australia are being considered. If built in South Africa, the Karoo Array Telescope (MeerKat) will play a central role, with other dishes spiraling outward across Africa.

Several other very large or extremely large telescopes are operational, under construction, or in development. Some notable ones are the Expanded Very Large Array (EVLA), the European Extremely Large Telescope (E-ELT), the Multi Mirror Telescope (MMT), South Africa's Large Telescope (SALT), Gran Telescopio Canaras (GTC), and the Allen Telescope Array (ATA, designed for the Search for Extraterrestrial Intelligence, SETI).

KNOWLEDGE GAINED

In the past, ground-based telescopes have helped scientists learn more about our solar system. Among the most important ways are the following: gathering of basic data on planetary celestial mechanics; discovery and orbit determination of comets and asteroids; mapping of galactic X-ray and infrared sources; discovery of quasars and radio galaxies; detection of cosmic background radiation; and the search for extrasolar planets.

In order to determine the continuously changing orbits of the planets well enough to make interplanetary spaceflight possible, it is important to map planetary positions very accurately. Laser ranging of the Moon, for example, allows astronomers with large ground-based telescopes to measure its distance to within a few centimeters. Radar-ranging measurements of the planets, especially Venus, Mercury, and Mars, have led to very precise determinations of their positions.

Spacecraft exploration of comets depends on ground-based telescopes, to discover comets in the first place and then to monitor them in their somewhat unpredictable paths near the Sun and Earth. Exploration of asteroids by spacecraft similarly depends on ground-based telescopes for tactical support.

Most of the cosmic objects found at X-ray and infrared wavelengths by orbiting detectors would be unexplained were it not possible to study them at other wavelengths from the ground. For example, two of the objects thought to be black holes, LMC X-1 and Cygnus X-1, would merely be mysterious, unidentified X-ray sources if it had not been possible, using large ground-based telescopes, to discover that each is a binary star, with a normal star in orbit around a dark, massive object.

The Keck telescopes in Hawaii have been used to study seasonal variations on Uranus. In 2007, astronomers were able to photograph the change of seasons on Uranus for the first time. Still in its early stages, this study found significant changes in some cloud features that had previously appeared to be unchanging. Scientists were also able to calculate wind speeds on the planet more extensively, up to 901 kilometers an hour. That year astronomers were also able to use the Keck telescopes to look at Uranus's ring system edge-on. They found that Uranus's dusty rings have changed since Voyager 2 visited the system in 1986. This ring crossing was also observed using the Very Large Telescope (VLT) in Chile, and the Palomar Observatory in Southern California. Such observations would not be possible, or economically feasible, without ground-based instruments.

As ground-based telescopes increase in size, they also increase in power. The newer generations of VLTs (very large telescopes) and ELTs (extremely large telescopes) allow scientists to study objects at greater distances and therefore look back in time, since their light takes years (light-years) to reach us.

CONTEXT

Astronomy seeks a better understanding of cosmology. Ground-based telescopes have provided the basic list of objects and phenomena (including quasars, radio galaxies, and cosmic background radiation) that allow astronomers to penetrate to the edge of the universe and the beginning of time. New generations of ground-based telescopes are designed to push farther into deep space and to determine the true story of how the universe came about. They study distant galaxies, stars, and planetary systems to learn about the formation and evolution of our own solar system.

Despite the advent of space-based observatories, ground-based observations will continue to play a major role in astronomy so long as the risks and costs associated with space-based telescopes continue to be high. Ground-based observations will therefore continue to provide data and images complementary to space-based observations, which together will aid astronomers in understanding the nature and evolution of the physical universe.

Jennifer L. Campbell

FURTHER READING

Anderson, Geoff. *The Telescope: Its History, Technology, and Future*. Princeton, N.J.: Princeton University Press, 2007. Anderson does a good job of summarizing the history of telescopes from their earliest days to those currently under construction. Intended to give the general reader the basics behind telescope design and a notion of current advancements in the field.

Brunier, Serge, and Anne-Marie Lagrange. *Great Observatories of the World*. Buffalo, N.Y.: Firefly Books, 2005. A photographic tour of fifty-six of the world's most impressive observatories. Decribes how telescopes work, as well as new advances in ground-based telescopes and the higher-resolution photographs they produce. For the general reader.

Kenyon, Ian. *The Light Fantastic: A Modern Introduction to Classical and Quantum Optics*. New York: Oxford University Press, 2008. An introductory work on optics aimed at advanced undergraduate and first-year graduate students. Also gives the reader practical examples and real-world applications. Includes a separate chapter on telescopes.

Kirby-Smith, Henry T. *U.S. Observatories: A Directory and Travel Guide*. New York: Van Nostrand Reinhold, 1976. A complete description of U.S. observatories, with some history and an account of the types of research and the equipment available. Notes on public availability for some are included. Suitable for general readers.

Kloeppel, James E. *Realm of the Long Eyes: A Brief History of Kitt Peak National Observatory*. San Diego: Univelt, 1983. A well-illustrated account of the development of the optical observatory on Kitt Peak, Arizona. Offers a complete history of the arguments for a national ground-based observatory and of the process of its establishment as one of the world's most productive scientific installations.

Krisciunas, K. *Astronomical Centers of the World*. Cambridge, England: Cambridge University Press, 1988. Provides thorough coverage of the world's major astronomical centers, including important historical installations. Includes material on advanced concepts, observatories of the future, and space observatories. Features useful tables and an index. Suitable for scientifically oriented general readers.

Kuiper, Gerard P., and Barbara M. Middlehurst, eds. *Telescopes: Stars and Stellar Systems*. Chicago: University of Chicago Press, 1978. A compendium of fairly technical chapters on various aspects of telescope design, with an emphasis on traditional large reflecting telescopes with equatorial mounts. Contains many tables, photographs, and graphs. For college-level readers.

Tucker, Wallace, and Karen Tucker. *The Cosmic Inquirers: Modern Telescopes and Their Makers*. Cambridge, Mass.: Harvard University Press, 1986. An excellent book about ground-based telescopes, including radio and space telescopes, and about the astronomers who brought them into existence. Includes narrative accounts and biographical data based on interviews with key people. For general readers.

Zirker, J. B. *An Acre of Glass: A History and Forecast of the Telescope*. Baltimore: Johns Hopkins University Press, 2005. Zirker's work provides a detailed study of telescope construction from the past to the future. It also explains their scientific significance. Answers any questions the general reader might have about telescopes.

See also: Archaeoastronomy; Coordinate Systems; Earth System Science; Gravity Measurement; Hertzsprung-Russell Diagram; Infrared Astronomy; Neutrino Astronomy; Optical Astronomy; Radio Astronomy; Telescopes: Space-Based; Ultraviolet Astronomy; X-Ray and Gamma-Ray Astronomy.

Telescopes: Space-Based

Category: Scientific Methods

Since the 1960's, robotic observatories placed in space have become a key part of the overall effort by astronomers to understand the solar system and universe. Able to view the universe without the interfering effects of Earth's atmosphere, these space telescopes have revealed sights and wonders that were previously unimagined.

OVERVIEW

With the exception of relatively few objects found within the solar system, objects of study outside our planet are too far away to visit and sample. We cannot simply travel to distant stars, nebulae, and galaxies to learn about them. Thus, if humans wish to study these distant objects, they must find another way to do so.

Electromagnetic Radiation. The primary method of learning about distant objects is by observing various kinds of electromagnetic radiation. One must bear in mind that since light travels at a finite speed (300,000 kilometers per second, or 186,000 miles per second), objects can be studied only as they were when they emitted the light we observe. Thus the telescope, whether in space or on Earth, can be thought of as something like a time machine. For example, if an object is determined to be 10 billion light-years away (a light-year being the distance traveled by light in vacuum in one year, that is, 9.467 trillion kilometers), the telescope provides a view of what that object looked like ten billion years ago. Depending on how distant objects are, they probably are radically different today from what they were in the past. For example, considering the presently accepted value for the age of the universe, a galaxy 12 billion light-years away is observed as it was only a little less than 2 billion years after the big bang.

Sir Isaac Newton demonstrated that visible white light is actually composed of an infinite sequence of wavelengths from red to violet in color. If one disperses white light with a prism or diffraction grating, then a rainbow band of colors called a spectrum is seen. Visible light is actually only a very small part of the total electromagnetic spectrum; the rest of it is made of other radiation, "invisible" light. If humans could see the other parts of the spectrum, they would find in turn beyond the violet end of the spectrum, at ever shorter wavelengths, ultraviolet, X-ray, and gamma-ray radiation. Beyond the red end, at ever longer wavelengths, are infrared, microwave, and radio radiation. Radio radiation includes frequencies used for broadcast, radar, and television signals. All of these different kinds of "light" are electromagnetic waves, or electromagnetic radiation, with specific frequencies and wavelengths. Electromagnetic radiation is such that all waves share the property that the product of their wavelength and frequency is the speed of light, approximately 300,000,000 meters per second (yielding the per-year results that define the light-year described previously). When taken together, these waves of spatially and temporally oscillating crossed electric and magnetic fields (or streams of photons) comprise the electromagnetic spectrum.

The instrument commonly used to collect light is the telescope. The telescope can be defined as a "light bucket," the function of which is to gather as much light as possible from a given region of the sky and direct it to a focal point. Today telescopes exist that study every part of the electromagnetic spectrum. The first telescopes, optical telescopes, collected light only in the visible region. These are still the most common type of telescopes today, especially on Earth's surface.

The opening in the telescope permitting entry of light is called the aperture, and its size serves two functions. The larger the aperture, the more light can enter, enabling one to see fainter objects. Larger apertures also result in greater resolution. "Resolution" pertains to the ability to "see" something in sharp focus. The smaller an object, the more resolution one needs to see it clearly. However, if greater resolution were the sole requirement for viewing faint objects, it would simply be necessary to make larger and larger telescopes in order to see them. The biggest problem facing astronomers, however, is Earth's atmosphere. Because of its chemical and physical composition, most electromagnetic frequencies are blocked out. Only certain frequencies have the ability to penetrate the atmosphere all the way to Earth's surface. One obvious frequency band is visible light. Another is radio waves. Other frequencies are absorbed or reflected by the atoms and molecules in the atmosphere, by water vapor, or by the natural or industrial dust and pollution in the air. If it were possible to see in any of these absorbed frequencies, such as ultraviolet light or X rays, the air would be opaque, like a wall. However, when instruments can be positioned beyond the atmosphere, rising above all the components of the atmosphere that block out invisible light, they are able to "see" in these invisible frequencies.

For these reasons, telescopes have historically been built in high places such as mountaintops. Telescopes have also sent aloft in airplanes, balloons, and rockets for short-term observations. However, it was not until the later half of the twentieth century, with the advent of the space age, that scientists were able to add telescopes and other detectors of electromagnetic radiation to orbiting spacecraft and interplanetary space probes in order to capture imaging data beyond Earth's atmosphere for long periods of time. As a result, new areas of astronomy—such as infrared, X-ray, and gamma-ray astronomy—blossomed. For example, infrared measurements from space are important to astronomers because infrared radiation is related to heat, and star-forming regions of the universe release large amounts of infrared radiation that is not detectable from Earth's surface. Similarly, high-energy astrophysics is studied by X-ray and gamma-ray observatories placed in Earth orbit.

Orbiting Solar Observatories. The idea of placing telescopes into orbit is an old one. The early space theorist Hermann Oberth wrote about this in the 1920's. In 1946 the astronomer Lyman Spitzer, Jr., composed a report titled *Astronomical Advantages of an Extra-Terrestrial Observatory*. Under the direction of its first chief astronomer, Nancy Roman, the National Aeronautics and Space Administration (NASA) prepared for the development of space-based observatories, an idea that was not initially welcomed by large portions of the astronomical community for fear that money would be diverted from the construction of ever larger ground-based observatories. NASA launched eight Orbiting Solar Observatories (OSOs) between 1962 and 1971, designed specifically to make observations of the Sun in visible, ultraviolet, and X-ray wavelengths. Seven of the eight OSOs were successful and provided data that added to our understanding of the eleven-year sunspot cycle, as well as characterizing the solar corona. An OSO produced the first full-duration image of the solar corona. Some of the OSO instruments were test beds for later space-based X-ray telescopes, such as those incorporated into the Skylab space station's Apollo Telescope Mount. OSO studies also investigated cosmic rays, neutron emissions, and extrasolar X-ray sources.

Orbiting Astronomical Observatories. The OSOs were followed by a sequence of Orbiting Astronomical Observatories (OAOs). The first OAO was launched on April 8, 1966. However, its pointing system failed when activated and caused the onboard batteries to explode. A successful OAO-2 was launched in December, 1968, making observations for four and a half years in the infrared, ultraviolet, X-ray, and gamma-ray portions of the electromagnetic spectrum. This advancement was followed by several other very successful orbiting observatories, including OAO-3 in 1972, the International Ultraviolet Explorer (IUE) in 1978, and the European Space Agency's (ESA's) Infrared Astronomical Satellite (IRAS) in 1983.

Granat. The Russians operated an X-ray/gamma-ray observatory named Granat from December, 1989, to November, 1998. It probed the center of the Milky Way, recorded spectral and temporal variability in black holes, and detected 511-kiloelectron volt (keV) X rays produced by electron-positron annihilation.

Extreme Ultraviolet Explorer. NASA's Extreme Ultraviolet Explorer (EUVE) was launched on June 7, 1992. This observatory was designed to detect and record radiation ranging in wavelength from 7 to 76 nanometers, a portion of the electromagnetic spectrum that had not been investigated by many space-based instruments previously. The telescope was outfitted with an imaging microchannel plate detector and three spectrometers available at its focal plane.

EUVE's mission continued through January 31, 2001.

The Great Observatories. In the 1980's NASA announced its Great Observatories Program, the purpose of which was to place in orbit a series of telescopes to observe across the electromagnetic spectrum. Observatories under this program include the Hubble Space Telescope (HST), which was deployed in orbit from the space shuttle *Discovery* (mission STS-31) on April 25, 1990; the Compton Gamma Ray Observatory (GRO), which was deployed in orbit from the space shuttle *Atlantis* (mission STS-37) on April 8, 1991; the Advanced X-Ray Astrophysics Facility (AXAF), which was deployed from the space shuttle *Columbia* (mission STS-93) on July 23, 1999; and the Space Infrared Telescope Facility (SIRTF), launched on August 25, 2003, by an expendable Delta II booster. After deployment, GRO was commissioned as the Compton Gamma Ray Observatory in honor of the American physicist Arthur Compton; AXAF was commissioned as the Chandra X-Ray Observatory (CXO) in honor of Indian astrophysicist Subrahmanyan Chandrasekhar; and SIRTF was commissioned as the Spitzer Space Telescope (SST) in honor of the American astronomer Lyman Spitzer, Jr.

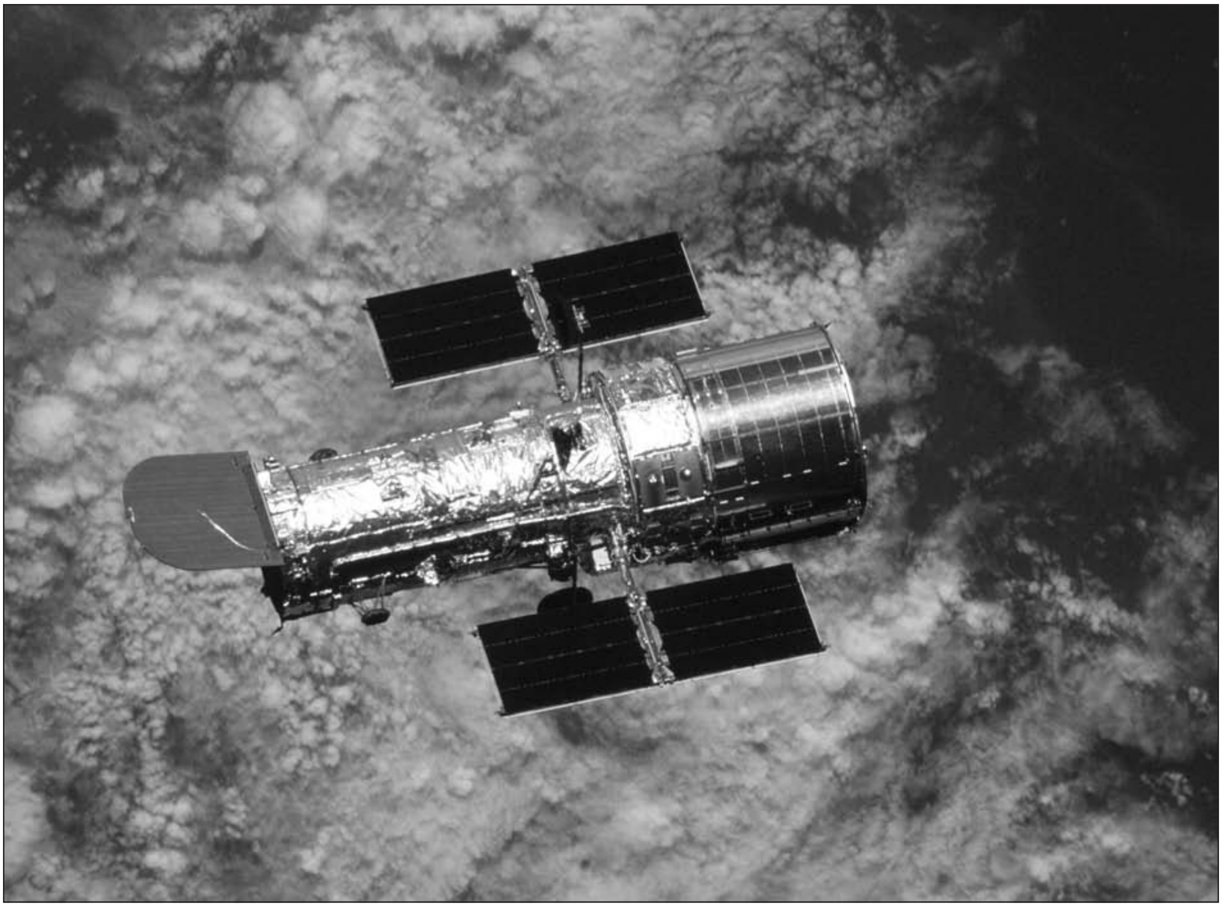
Many of the observatories described thus far represented revolutionary steps in expanding humanity's investigations of the universe. Many opened up entirely new windows on the universe, since the types of radiation they detected could not penetrate Earth's atmosphere to reach ground-based telescopes. In order to gain a clear picture of the physical processes in stars, galaxies, and exotic objects such as quasars, active galactic nuclei, gamma-ray bursters, and black holes, it is necessary to examine those objects simultaneously across the different regions of electromagnetic spectrum. That indeed was the overriding aim of NASA's Great Observatories Program: to construct several cutting-edge telescopes that, once in space, would provide means to observe across the spectrum, from gamma rays to infrared radiation. No radio telescope in space was included in this project, since the majority of the radio spectrum is observable from ground-based facilities. Nevertheless, for reasons of having an electromagnetic quiet zone away from earthly sources, many astronomers have long dreamed of a radio telescope on the far side of the Moon.

Hubble Space Telescope. The crown jewel of all these Great Observatories is certainly the Hubble Space Telescope (HST); it has had the greatest appeal to the public. Hubble was launched aboard the space shuttle *Discovery* into the highest attainable space shuttle orbit, at 615 kilometers above Earth's surface, with an orbital inclination of 28.5°. After requiring a heroic on-orbit repair by space shuttle mission STS-61 astronauts, Hubble began producing a steady stream of impressive images that fascinated the public and rewrote astronomy textbooks. For example, by early 1997 Hubble had taken more than 100,000 images of various objects in the solar system and universe, amazing scientists and the general public alike with the incredible sights it had captured. Hubble is expected to continue providing new insights about the universe well into the second decade of the twenty-first century.

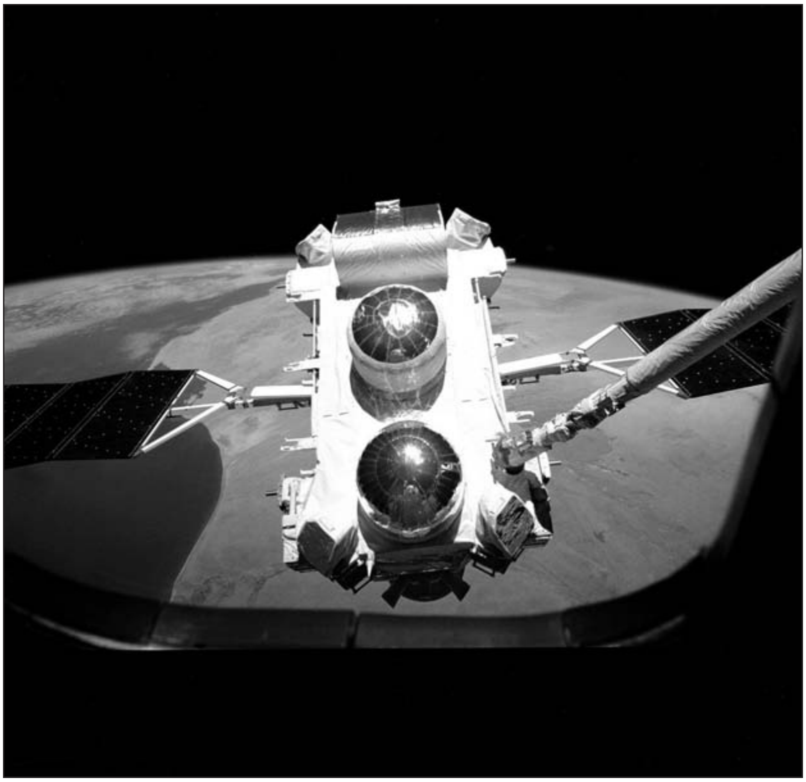
A telescope's ability to see is controlled by the size of its aperture, but it is also controlled by the medium through which the telescope must look (or what the radiation must travel through) to capture its images. Even though visible light can penetrate the air, the atmosphere has several detrimental effects on a telescope's ability to see. These effects include twinkling, which is caused by movement in the air; weather; the moisture content of the air; dust and air pollution; the natural glowing of the atmosphere, known as airglow; and light pollution caused by city lights. Because of all these factors, observatories are built on remote mountaintops. The largest optical telescopes in the world are the two 10-meter (393.7-inch) Keck telescopes on Mauna Kea in Hawaii and the 11-meter (433-inch) Hobby-Ebberly telescope at the McDonald Observatory on Mount Locke, Texas. Even these instruments, however, contend with atmospheric effects. By placing a telescope in orbit, all of these atmospheric factors are eliminated, making it possible to obtain much clearer images. The Hubble Space Telescope has a primary mirror of only 2.4 meters (94.5 inches), yet it can see many times better than either of these two earthbound telescopes.

Hubble was designed to provide the clearest view ever of the cosmos, providing astronomers with vision that was ten times clearer and fifty times more sensitive than that of the best ground-based telescopes. This provides astronomers with the largest boost in viewing capability since Galileo first used a telescope to view the sky in 1610. Hubble is 13.3 meters long, 4.3 meters in diameter, and 12 meters across with its solar array deployed. It weighs 11,200 kilograms (12.3 tons). The 2.4-meter primary mirror is a Cassegrain mirror. That is, the light reflected by the primary mirror is then reflected by a secondary mirror, after which the light travels through a hole in the primary mirror to the instruments positioned behind it.

The primary mirror was manufactured with greater precision than anything previously. Yet, shortly after the telescope was deployed in 1990, it became apparent that there was a problem. Although the mirror was extremely smooth, it was not shaped correctly, exhibiting what is known as a spherical aberration—that is, the curvature of the mirror was slightly rounder than the parabolic shape required. As a result, the images were blurry. However, this problem was rectified by means of correcting lenses and mirrors known as the Corrective Optics Space Telescope Axial Replacement (COSTAR). COSTAR, along with several new gyroscopes and instruments, was installed by the STS-61 space shuttle servicing mission in December, 1993. A second servicing mission (STS-82) occurred in February, 1997. Additional servicing missions were conducted in late 1999 (STS-103) and in 2002 (STS-109). These servicing missions allowed new instruments to be emplaced within the telescope and accom-



The Hubble Space Telescope in orbit around Earth, 2002. (NASA)



The Compton Gamma-Ray Observatory at its deployment by the space shuttle Atlantis *in 1991.* (NASA)

plished maintenance and repair work to keep Hubble operational for its designed fifteen-year life span, until 2005. In May, 2009, a shuttle-based repair mission, STS-125, was flown to keep the Hubble Space Telescope operational at least through 2011-2013, when the next-generation space telescope, the James E. Webb Space Telescope (JWST), was expected to supplant Hubble.

Compton Gamma Ray Observatory. The Compton Gamma Ray Observatory (CGRO) was deployed in a high orbit (450 kilometers above the Earth's surface), yet still underneath the Van Allen radiation belts to preclude interference with the observatory's detectors. Compton was outfitted with four instruments: the Burst and Transient Source Experiment (BATSE), the Oriented Scintillation Spectrometer Experiment (OSSE), the Imaging Compton Telescope (COMPTEL), and the Energetic Gamma Ray Experiment Telescope (EGRET). Together these scintillation detectors spanned six orders of magnitude of wavelengths in the gamma-ray portion of the electromagnetic spectrum, specifically from 20 keV to 30 gigaelectron volts (GeV) in energy.

When Compton was on the verge of losing attitude control, NASA decided to deorbit the heavy observatory in a controlled fashion so that any debris, such as portions of its very massive detectors, would not impact populated areas. Compton was deorbited on June 4, 2000, and safely rained debris over open regions of the Pacific Ocean.

Infrared Space Observatory. ESA followed up its highly successful IRAS cataloging mission with the Infrared Space Observatory (ISO). After launch on November 17, 1995, ISO produced more than twenty-six thousand infrared observations before its liquid helium cryogenic coolant supply ran out after a bit more than twenty-eight months. ISO was outfitted with a high-resolution infrared camera (ISOCAM), an infrared photo-polarimeter (ISOPHOT), a shortwave spectrometer (SWS), and a long-wave spectrometer (LWS) in order to observe wavelengths between 2.4 and 240 microns. Among ISO's achievements were the finding of planet formation around old stars nearing their end of life, the detection of dust between galaxies, finding a number of new chemical compounds in interstellar gas clouds, locating protoplanetary disks around stars in early stages of formation, and determining chemical abundances in planets within this solar system.

Far Ultraviolet Spectroscopic Explorer. NASA launched the Far Ultraviolet Spectroscopic Explorer (FUSE), an Explorer-class satellite (specifically, Explorer 77), on June 24, 1999, atop a Delta II booster. FUSE, an observatory in the agency's Origins program, was placed in a low Earth orbit. The spacecraft was designed to last at least three years. Scientists at the Johns Hopkins University used the spacecraft's Wolter-type grazing incidence telescope and its spectrograph to observe far-ultraviolet emissions ranging from 90.5 to 119.5 nanometers. FUSE was designed specifically for determining the distribution of deuterium (the isotope of hydrogen having a nucleus consisting of a proton and a neutron) in the aftermath of the big bang. FUSE covered a region of the electromagnetic spectrum not detected by any of the members of the Great Observatory program and investigated astrophysical and cosmological questions that added to and complemented discoveries made by Hubble, Chandra, and other space-based telescopes.

FUSE suffered failures in its pointing system and for a time was on the verge of being shut down, despite the fact that its science hardware functioned perfectly. Engineers developed alternate means of pointing the telescope, and FUSE got a reprieve for more science. However, on July 12, 2007, the spacecraft lost its last reaction wheel. Efforts to restore FUSE to science operations failed, and two months later the telescope was abandoned.

Chandra X-Ray Observatory. Because X rays cannot be focused in the same manner as visible light, the design of the Chandra X-Ray Observatory required a quartet of nested pairs of mirrors in the shape of paraboloids and hyperboloids. At normal incidence (that is, perpendicular to the surface of the mirror), X rays are much more likely to be absorbed by than reflected off a mirror; hence, there the design for any X-ray telescope must allow for low grazing angles. Chandra's X-ray resolution was over a thousand times greater than any previous space-based X-ray detector. Placed in a highly elliptical orbit ranging from 10,000 to 140,000 kilometers, Chandra is capable of collecting data during as much as 85 percent of its sixty-five-hour orbital period. Chandra is outfitted with four instruments: an imaging spectrometer, a high-resolution camera, and high-energy and low-energy high-resolution transmission grating spectrometers.

Spitzer Space Telescope. The Spitzer Space Telescope was placed in an Earth-trailing heliocentric orbit in 2003. The telescope was designed to be cooled to 5.5 kelvins and last a minimum of 2.5 years. Outfitted with an infrared camera (the Infrared Array Camera), an infrared spectrometer (the Infrared Spectrograph), and far-infrared detection arrays (the Multiband Imaging Photometer for Spitzer), Spitzer was designed to cover infrared emissions from 3 to 180 microns. Spitzer's supply of liquid helium cryogenic coolant lasted longer than a minimal mission, and, in late 2008, it was calculated that Spitzer could last between another six months and a year, or perhaps longer.

Swift Gamma-Ray Burst Mission. The Swift Gamma-Ray Burst Mission, usually referred to simply as Swift, was launched on November 20, 2004, to study and identify the origin of gamma-ray bursts. Aptly named, Swift had the capability to reorient itself quickly in order to place an incoming gamma-ray flux centrally within the observatory's field of view, and thereby alert the astronomical community rapidly so that other telescopes and detectors could record the afterglow of a gamma-ray burster (GRB), even short-period ones.

Gamma-Ray Large Area Space Telescope. The Gamma-Ray Large Area Space Telescope (GLAST) was launched on June 11, 2008. It was a next step in gamma-ray astronomy after the loss of the Compton Observatory. GLAST was designed to investigate black holes, the high-speed jets of gas and energetic particles emitted from black holes, the physics of dark matter, solar flares, pulsars, cosmic-ray production, and gamma-ray bursts.

Kepler Observatory. The Kepler Observatory was designed to survey more than 100,000 stars in four years to detect transits of exoplanets across those stars' photospheres. Such a crossing would result in a slight diminishment of the star's light curve, and additions to the stellar spectrum of absorption lines from the planet's atmosphere, if it were to have one. Planned for launch in early 2009, Kepler's primary objective was to detect terrestrial planets within habitable zones around a wide variety of stars. The location of habitable zones is determined by a star's spectral class and luminosity. For a G2-class star such as the Sun, the habitability zone—where life as humanity knows it is permissible due to the possibility of water existing in liquid form—extends from 0.72 astronomical units (AU) out to 1.5 AU, essentially from Venus to Mars. However, for a red dwarf star, the habitability zone would be closer. The chance of detecting an Earth-class planet transiting a star is a complicated combination of star spectral class, the possibility of planets forming about that particular star, and the planet's mean distance from the star.

For example, for the star Tau Ceti (spectral class G8), if a terrestrial planet similar to Earth were located 1 AU away from the star, the possibility of a transit being observed is approximately 1 in 210. By the sheer number of survey stars Kepler would examine, the chances of finding evidence of terrestrial planets would be high. It was estimated that if Kepler were able to examine its planned 100,000 stars, then 480 or more terrestrial planets should be observed by this transit method.

James E. Webb Space Telescope. Originally called the Next Generation Space Telescope (NGST), the James E. Webb Space Telescope (JWST, named for James E. Webb, who served as NASA Administrator from 1961 to 1968) has often been described as the replacement for the Hubble Space Telescope. This is not strictly accurate, as JWST is designed as an exclusively infrared observatory, to be placed in a position where the Sun's gravitational attraction on the telescope is balanced precisely by that of the Earth. This location is called the L2 Lagrangian point. It is located 1.5 million kilometers from Earth.

JWST is designed to be constructed of eighteen hexagonal mirror segments that must deploy from a folded configuration at launch. If successfully deployed, JWST will have a light-collecting area six times greater than Hubble's. Instruments incorporated into JWST's design included a near-infrared camera, a near-infrared spectrograph, a mid-infrared instrument, and fine guidance sensors. These will provide coverage of the infrared between wavelengths of 0.6 to 28 microns. Launch of JWST is planned for sometime in the 2010's. Should Hubble still be available at the time, early studies conducted by JWST could be coordinated with those supplied by Hubble. Among JWST's primary objectives are an investigation of the earliest epoch of star and galaxy formation, observation of evolutionary processes in galaxies, examination of star-forming regions and developing planetary systems, and searches for planetary systems that may have materials and conditions necessary for the evolution of life.

KNOWLEDGE GAINED

Findings from OAOs. Among the achievements of the Orbiting Astronomical Observatories were the finding that halos of hydrogen surround comets, the discovery that novae may increase in ultraviolet brightness while diminishing in visible luminosity, the discovery of long-period pulsars, and the collection of several hundred high-resolution spectra of stars in X-ray and ultraviolet wavelengths.

Findings from IRAS. The Infrared Astronomical Satellite was the first fully dedicated infrared observatory in space. Its greatest achievement was the location of an enormous number of infrared sources to be studied in detail by later generations of infrared detection systems placed in space, such as ISO and Spitzer.

Findings from IUE. Among the scientific discoveries of the International Ultraviolet Explorer were auroras in Jupiter's atmosphere, sulfur in comets, stellar spots, the progenitor star for Supernova 1987A, the extent of active regions in Seyfert galaxy nuclei, and galactic halos. IUE produced the first light curves collected undisturbed for more than twenty-four hours, and it detected the first emissions of wavelengths less than 50 nanometers. More than 104,000 observations were made by IUE before its decommissioning.

Findings from Hubble. One of the primary goals for Hubble was to measure the brightness of certain kinds of stars, called Cepheid variables, in several distant galaxies, making it possible to determine how far away these galaxies are. The importance of this is related to a discovery by the American astronomer Edwin Hubble, the man after whom the Space Telescope is named. Hubble discovered that the universe is expanding and that the distance to an object is proportional to just how far away the object is. This proportionality is known as Hubble's constant. If one finds a reliable value for Hubble's constant, it is possible to determine just how far away a galaxy is. It is also possible to determine how old the universe is. The problem is finding a reliable value. Using measurements made with the Hubble after the installation of COSTAR, astronomers were able to determine a definitive value for this constant. In 2008 the accepted value for the age of the universe was determined to be 13.7 billion years, and for the Hubble constant, 77 kilometers per second per megaparsec.

Hubble was not necessarily expected to make significantly new discoveries, because it observes in the same visible light in which observations from the Earth's surface are made. Rather, it was expected to see with unequaled clarity and resolution the objects previously observed. Yet, it has still succeeded in making previously unsuspected discoveries, mainly because it is able to see much fainter objects than could ever be seen using ground-based telescopes. One such discovery occurred during a ten-day exposure of a section of the sky that was previously thought to be empty and was found by Hubble to be filled with distant galaxies. Two such Hubble deep field surveys were eventually made. The other Great Observatories in orbit made radically new discoveries from the very beginning, because they observe in light frequencies in which humans were previously unable to observe due to atmospheric influences.

While it is the mirror that actually gathers the light, it is the instruments on the Hubble Space Telescope that provide its tremendously exciting scientific advances. One of these instruments is the Wide Field/Planetary Camera-2 (WF/PC-2). WF/PC-2 serves as the primary imaging camera. This instrument has provided us with the startling images of towers of interstellar matter and other phenomena that have become so familiar.

Another instrument is the Space Telescope Imaging Spectrograph (STIS), which gives the Hubble unique and powerful spectroscopic capabilities. A spectrograph separates light gathered by the telescope into its spectral components so that the composition, temperature, motion, and other chemical and physical properties of astronomical objects can be analyzed. STIS's two-dimensional detectors allow the instrument to gather thirty times more spectral data and five hundred times more spatial data than existing spectrographs on the Hubble that observe one location at a time. The STIS is a particularly powerful tool for studying supermassive black holes. STIS searches for massive black holes by studying the star and gas dynamics around galactic centers. It also measures the distribution of matter in the universe by studying quasar absorption lines, using its high sensitivity and spatial resolution to examine star formation in distant galaxies and performing spectroscopic mapping of solar-system objects.

Another instrument is the Near Infrared Camera and Multi-Object Spectrometer (NICMOS), which promises to gather valuable new information on the dusty centers of galaxies and the formation of stars and planets. Consisting of three cameras, NICMOS is able to perform infrared imaging and spectroscopic observations of astronomical targets. Because these detectors perform more efficiently than previous infrared detectors, NICMOS has given astronomers their first clear view of the universe at near-infrared wavelengths between 0.8 and 2.5 micrometers. These views in the infrared are important because expansion of the universe has shifted the light from very distant objects toward longer red and infrared wavelengths. Hence, NICMOS's near-infrared capabilities provide views of objects too distant for research by Hubble's optical and ultraviolet instruments.

Astronauts installed the Advanced Camera for Surveys (ACS) inside Hubble in 2002, replacing the Faint Object Camera. ACS was designed to observe from far-ultraviolet to visible wavelengths, and therefore to collect images that would shed light on some of the earliest galaxies. ACS possesses a field of view twice that of the Wide Field Planetary Camera 2, giving it the capability to accomplish very broad surveys of the early universe. ACS lost performance in its charge-coupled devices in early 2007 and was slated for repair during the STS-125 final Hubble servicing mission.

Included as a part of a final planned shuttle service mission to the Hubble (the shuttle flight designated STS-125) was installation of the Cosmic Origins Spectrograph (COS), designed to perform spectroscopy between 115 and 320 nanometers in ultraviolet wavelength. According to the Space Telescope Science Institute, this instrument's principal research would center on investigating the large-scale structure of the universe, the early formation of galaxies and their subsequent evolution, formation of stellar and planetary systems, and the interstellar medium. Several spacewalks would be required on STS-125 to achieve a life extension for Hubble. The mission was set up to install six new gyrodynes, six new solar batteries, the COS, and the Wide Field Camera 3, a new cooling system, and a replacement Fine Guidance Sensor. The astronauts would also repair both the thermal insulation and STIS, in order to leave Hubble in an almost new configuration that would last at least five years.

Among the numerous major discoveries from ACS research was the collection of gravitational lensing data sufficient to determine the mass of seventy distant galaxies. Gravitational lensing is the process whereby massive objects bend and focus light passing near them so that a telescope will see a ring of identical images, something referred to as an Einstein ring. The amount of lensing present in an image is determined by the mass of the lensing objects, such as these seventy galaxies, which divert the light and form the Einstein ring.

Findings from Compton. Among the Compton Gamma Ray Observatory's impressive achievements over its nine years in orbit were the amassing of an all-sky survey of emissions over 100 megaelectron volts (MeV). Compton identified 271 individual sources, compiling an all-sky map of gamma-ray emissions from decays of the radioactive isotope Al^{26} of aluminum. Compton also detected gamma-ray burster 990123, the brightest object recorded to that time, along with gamma-ray emissions from the tops of terrestrial thunderstorms.

Findings from EUVE. The Extreme Ultraviolet Explorer's mission consisted of two phases. During the first, which lasted only half a year, the telescope was used in imaging mode to generate a full-sky atlas of sources of extreme ultraviolet radiation. The telescope's second phase of investigation then involved pointed studies of individual sources using the spectroscopic capabilities of EUVE. Among EUVE's discoveries and observations with regard to the solar system were extreme ultraviolet/X-ray emissions in Comet P/Encke, changes in interplanetary helium wind, day glow at Venus, mechanisms whereby solar flares promote coronal heating, and images of the full Moon in extreme ultraviolet wavelengths.

Findings from Chandra. The Chandra X-Ray Observatory's contributions to X-ray astronomy include the first detection of X-ray emissions from the Milky Way's supermassive black hole located at the galactic center; detecting a mid-mass-range black hole in galaxy M82; providing data which might be evidence of a star composed of matter collapsed to quarks (a quark star); demonstrating that virtually all main sequence stars also emit somewhat in the X-ray region of the electromagnetic spectrum; helping to determine the Hubble constant (and hence the age of the universe); and collecting evidence of dark matter involved in supercluster collisions.

In 2008 Chandra and the Very Large Array of Earth-based radio telescopes located near Socorro, New Mexico, collaborated to observe the most recent supernova in the Milky Way. This supernova had occurred only 140 years earlier. Prior to this discovery, the most recently noted supernova in the galaxy was Cassiopeia A. The data on this young supernova allowed astronomers to investigate the stellar explosion process and creation of a central neutron star or black hole. This discovery indicated how useful coordination of observations in both X-ray and radio emissions as well as coordination of space-based and Earth-based telescope facilities can be.

Another unique way Chandra was used involved developing a new method for determining the mass of supermassive black holes, such as those found at the centers of many galaxies. Looking for X-ray emissions from hot gas in the central region of the elliptical galaxy NGC 4649, Chandra was able to determine the peak temperature of that hot gas. The temperature of the hot gas is determined by the gravitational compression of the gas from the black hole. That, in turn, is dependent on the black hole's mass. This method was checked against earlier methods of "weighing" black holes. The results from Chandra for NGC 4649 agreed nicely.

Findings from Spitzer. The Spitzer Space Telescope greatly expanded on previous infrared research produced by IRAS and ISO, and collaborated with NASA's other Great Observatories for coordinated studies of important objects and also deep-sky field surveys. Among Spitzer's major accomplishments are the first direct detection of light from the hot-Jupiter extrasolar planets HD 209458b and TrES-1; detection of the youngest star ever found; determination that the Milky Way galaxy's core has a bar structure; capturing the glow from stars formed only 100 million years or so after the big bang; and the first determination of an extrasolar planet's atmospheric temperature (that of HD 189733b).

Spitzer was used to study Messier 101, otherwise known as the Pinwheel galaxy, detailing the infrared signature of the galaxy's spiral arms and central region. Data revealed that polycyclic aromatic hydrocarbons were present throughout much of the galaxy but disappeared in the outer region. This suggested that hydrocarbons, and hence organic materials, had a threshold as one looked far away from the central region of this galaxy.

In 2008 Spitzer observed a pair of young stars, DR Tau and AS 205A, located 457 and 391 light-years from Earth, respectively, particularly concentrating on their protoplanetary disks using the telescope's spectrographs. Both stars' disks displayed emission lines of water vapor within their innermost portions. This discovery, although the amount of water found was still considerably less than the amount of water found in Earth's oceans, indicated that water can be available and abundant in the inner regions of forming solar systems.

Findings from Swift. The Swift Gamma-Ray Burst Mission located a gamma-ray burster (GRB) with a burst duration of only 0.05 second. Swift also found the brightest object ever seen, GRB 080319B, located 7.5 billion light-years from Earth with a luminosity 2.5 million times that of the most brilliant supernova seen previously.

Findings from European and Russian Observatories. Although this essay has primarily concentrated on NASA observatories in space, the Europeans and Russians have placed a number of important and productive space-based telescopes in orbit, including the previously discussed IRAS and ISO. Also important are XMM-Newton, an X-ray observatory; the Salyut 6 space station's KRT-10 radio telescope; and COROT. These are only a few of the international astrophysical facilities in space.

One of the many discoveries made by the European Space Agency's XMM-Newton X-ray observatory was the finding of an exploding star in the Milky Way that had been previously missed. This object was once so bright it could have been seen by the naked eye during the period of its initial explosion and subsequent nova phase. XMM-Newton accidentally found this nova on October 9, 2007, as it slewed from one planned target of opportunity to the next planned observation. The discovery reminded astronomers that serendipity in space-based astrophysics is just as important as it has been during the history of earthbound astronomical observations.

COROT was launched in December, 2006, to survey stars and search for extrasolar planets using a transit method. In this type of investigation, light from a star is seen to diminish by a very small but measurable amount when a planet passes in front of that star. Absorption of light by the planet's atmosphere reveals something about the nature of that atmosphere. By mid-2008, COROT had surveyed more than fifty thousand stars. A discovery in 2008 was particularly intriguing. COROT detected the presence of an extrasolar planet the size of Jupiter orbiting a star similar in mass to the Sun. Located close to this star, the Jupiter-class exoplanet completes an orbit in just 9.2 days. The star rotates at exactly the same rate, thereby synchronizing the planet and star.

CONTEXT

Earthbound observatories have not been made obsolete by space-based telescopes. The former continue to be needed in order to perform some of the most important investigations in contemporary astrophysics. Space-based telescopes, however, have opened the eyes of scientists to a surprising, wondrous, energetic, and violent universe that could not have been envisioned based solely on earthbound observations. Our view of the solar system has been dramatically changed by widening the portion of the electromagnetic spectrum to which astronomers gained access once observatories were operational in space. One is reminded of the poetic verse of T. S. Eliot: "We shall never cease from exploration, and the end of all our exploring will be to arrive where we started and know the place for the first time."

Among the best examples of space-based telescopes that have altered the general public's view of the universe, and the utility of studying it, is the Hubble Space Telescope. Hubble's dramatic images of towering columns of interstellar matter and other phenomena have replaced the common notion of a sterile and relatively empty universe with a new understanding of the endless variety and dynamism of the cosmos. Besides adding to our rapidly advancing scientific understanding of the universe, Hubble has directly contributed to the health, safety, and quality of people's lives through a variety of technological spin-offs. For example, a nonsurgical breast biopsy technique using a device originally developed for Hubble's Imaging Spectrograph is now saving women pain, scarring, radiation exposure, time, and money. Called stereotactic automated large-core needle biopsy, this technique enables a doctor to locate a suspicious lump precisely and use a needle instead of incisional surgery to remove tissue for pathology. This precise procedure has been rendered possible because of a key improvement in digital imaging technology known as the charge-coupled device, or CCD. In addition to such practical applications, perhaps Hubble's greatest contribution has been in the realm of education, where it has been extremely successful in generating enthusiasm for astronomy among both students and the public alike.

Christopher Keating and David G. Fisher

FURTHER READING

Bely, Pierre, ed. *The Design and Construction of Large Optical Telescopes*. New York: Springer, 2003. An engineering history of the design and construction of large telescopes, including the Hubble Space Telescope. Explains the astronomical capabilities of Hubble.

Chaisson, Eric. *The Hubble Wars*. New York: HarperCollins, 1994. A well-written account of the inside story on the development and deployment of the Hubble Space Telescope, written by someone who was there. Careful attention is paid to explaining the scientific issues in terms the layperson can understand. Suitable for all levels.

Christensen, Lars Lindberg, Robert A. Fosbury, and M. Kornmesser. *Hubble: Fifteen Years of Discovery*. New York: Springer, 2006. Details the accomplishments of Hubble over its first fifteen years. Explains the ability of Hubble to examine the universe from ultraviolet through the near-infrared portion of the electromagnetic spectrum.

Devorkin, David, and Robert W. Smith. *Hubble: Imaging Space and Time*. Washington, D.C.: National Geographic Society, 2008. An oversized picture book including a large number of high-quality color images taken by the Hubble over its first seventeen years of research. Published just in advance of the final servicing mission to the telescope. Thoroughly readable.

Kerrod, Robin, Carole Stott, and David S. Leckrone. *Hubble: The Mirror on the Universe*. New York: Firefly Books, 2007. An overview of the Hubble Space Telescope's achievements. Fully illustrated with color prints of spectacular Hubble images. Describes the astronomical significance of the greatest achievements of Hubble.

Tobias, Russell R., and David G. Fisher, eds. *USA in Space*. 3d ed. Pasadena, Calif.: Salem Press, 2006. This three-volume set consists of a series of well-written, in-depth articles on all American space missions, both crewed and uncrewed. A number of articles discuss individual space telescopes in detail. A major library reference designed for the nonspecialist audience, with tables, lists, and indexes, and more than 400 photos. Suitable for all levels.

Trumper, Joachin, and Gunther Hasinger. *The Universe in X Rays*. New York: Springer, 2008. Reviews the development of X-ray astronomy and its impact on advancements in astrophysics and cosmology. Covers results from ROSAT, RXTE, BeppoSax, Chandra, and XMM-Newton.

Voit, Mark. *Hubble Space Telescope: New Views of the Universe*. New York: Harry N. Abrams, 2000. Details the saga of Hubble from concept to the flawed telescope needing orbital repair to its status as a frontier astrophysics research facility. Includes large-format high-resolution images taken by Hubble which demonstrate the tremendous capability of this observatory to examine objects at all distances, from those within the solar system to those more than 12 billion light-years from Earth—nearly back to the beginning of the universe.

Weeks, T. C. *Very High Energy Gamma Ray Astronomy*. New York: Taylor & Francis, 2003. Covers gamma-ray astronomy through results from the Compton Gamma Ray Observatory. Designed for students of either theoretical or experimental high-energy astrophysics.

Wilkie, Tom, and Mark Rosselli. *Visions of Heaven: The Mystery of the Universe Revealed by the Hubble Space Telescope*. London: Hodder & Stoughton, 1998. Presents numerous color images of Hubble's greatest discoveries with narrative to explain their astrophysical significance.

Zimmerman, Robert. *The Universe in a Mirror: The Saga and the Visionaries Who Built It*. Princeton, N.J.: Princeton University Press, 2008. Focuses not only on the stunning images produced by Hubble but also on the history of astronomical investigations and the development, construction, launch, and on-orbit repair of Hubble. Not only a technological story, Zimmerman's book is a human saga of scientists and the political process of obtaining funding for such a big science project.

See also: Archaeoastronomy; Coordinate Systems; Earth System Science; Gravity Measurement; Hertzsprung-Russell Diagram; Infrared Astronomy; Neutrino Astronomy; Optical Astronomy; Radio Astronomy; Telescopes: Ground-Based; Ultraviolet Astronomy; X-Ray and Gamma-Ray Astronomy.

Terrestrial Planets

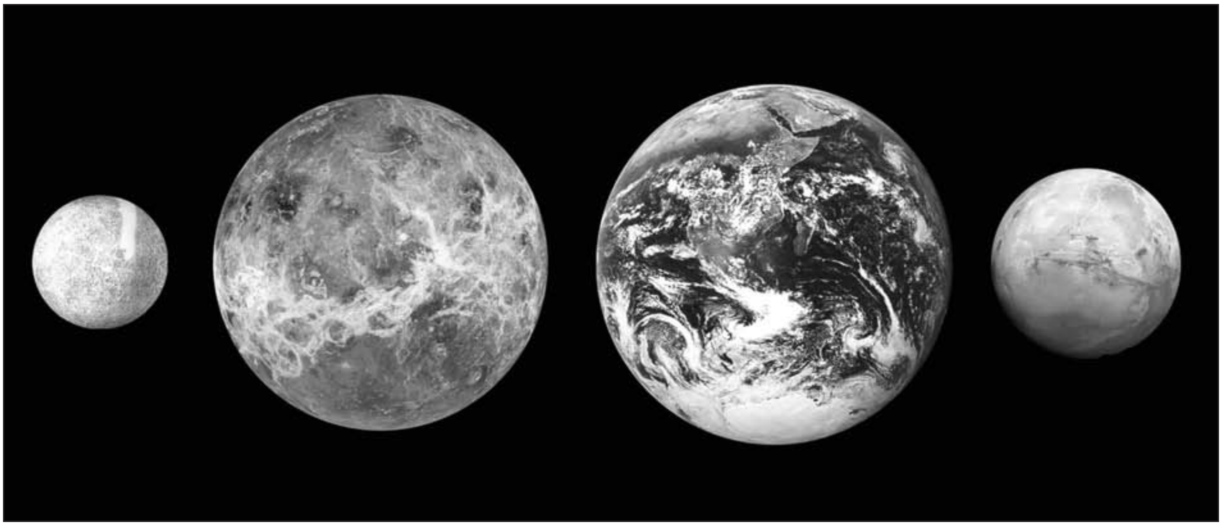
Category: Planets and Planetology

The terrestrial planets are the four inner planets of the solar system. Mercury, Venus, Earth, and Mars are respectively the closest to the farthest from the Sun. These planets are composed of mostly silicate minerals and a dense inner core composed mostly of iron.

OVERVIEW

In broad terms, the planets of the solar system can be divided into two basic types: terrestrial and gas giants. The terrestrial planets (Mercury, Venus, Earth, and Mars) are by and large rocky with varying degrees of atmospheric envelopes and different amounts of water, ranging from very little to a great deal. The gas giants (Jupiter, Saturn, Uranus, and Neptune), as the name implies, are larger than the terrestrial planets and have tremendous gas envelopes surrounding their cores, which remain hidden from direct observation.

Mercury is the smallest of the terrestrial planets, 4,880 kilometers in diameter. The density of Mercury is relatively high, 5.4 grams per milliliter, implying that it has a bigger iron core than that of the Earth. Mercury also has a weak magnetic field. The magnetic field has been difficult to explain, since models for heat flow predict that the iron core ought to be solid. The surface of Mercury contains many impact craters, so it looks much like the surface of the Earth's moon. These craters probably formed by many meteorites (planetesimals) bombarding Mercury early in the formation of the solar system. The sparse data on the composition of the surface of Mercury suggest that the iron content is fairly low, so it cannot be composed of volcanic rocks, common on the other terrestrial planets, called basalts. Instead, the low iron content suggests that the surface is composed mostly of material formed from the planetesimals hitting the surface. There is no evidence that volcanic activity occurred on Mercury. Mercury has almost no atmosphere. The trace of hydrogen or helium in the atmosphere is most likely derived from the Sun. Temperatures vary from about 623 kel-



The terrestrial planets, in this composite image, are shown together to feature their relative sizes. From left: Mercury, Venus, Earth, and Mars. (Lunar and Planetary Institute)

vins in sunlight down to 103 kelvins in darkness.

Venus has a diameter very similar to that of the Earth (12,100 kilometers). The density of Venus is high enough (5.2 grams per cubic centimeter) to suggest that it also has a dense iron-rich core. Predictions of the planet's heat flow suggest that this iron-rich core ought to be molten like that of the Earth, which on the Earth produces the strong magnetic field. Venus, however, has no magnetic field, and the reason for this lack has not been determined. The surface of Venus is sparsely cratered, with about 1,000 randomly distributed impact craters. The reason for the distribution of these few craters is that large lava flows of basalt were extruded over most of the surface about 500 million years ago, covering craters formed earlier, so only the most recent impact craters are exposed at the surface. Many other surface features have been observed on Venus. For example, highlands and lowlands form 20 percent of the surface, and midlands form 80 percent of the surface. Each of these areas includes features such as folds, fractures, lava flows, volcanoes, and features resulting from weathering. The composition of the rocks is similar to those of basalts found on the Earth. The atmosphere of Venus has about ninety times the pressure at the surface as that on Earth. The main constituent in the atmosphere is carbon dioxide (96 percent), with lesser amounts of nitrogen and sulfur dioxide. The surface temperature of Venus is 733 kelvins, a result of extreme greenhouse warming due to the carbon-dioxide-rich atmosphere, which absorbs infrared radiation. If the atmosphere were similar to that of Earth, the temperature would likely be about 348 kelvins.

The diameter of Earth is 12,756 kilometers. Earth has a density of 5.52 grams per cubic centimeter, and it has a dense, iron-rich core. The planet's inner core is believed to be solid, and the outer core is believed to be liquid. The moving liquid iron is believed to produce Earth's strong magnetic field. This magnetic field helps to protect life by deflecting many of the charged particles (the solar wind) coming from the Sun. Few impact craters can be found on Earth because its crust is continuously re-formed as a result of plate tectonics and weathering. Earth's crust consists of about twenty plates of varying sizes, averaging around 100 kilometers thick. These plates move laterally at less than about 10 centimeters per year from areas where material is being formed in volcanic mountain ranges to regions where one plate is being subducted under another plate. Other volcanism occurs in the subduction zones. In contrast to the other terrestrial planets, Earth contains abundant liquid water, which covers about 70 percent of the surface, mostly in the oceans. Certain organisms in the water have helped to remove much of the carbon dioxide from the atmosphere by forming carbonate rocks containing carbonate minerals. Removal of the carbon dioxide from Earth's atmosphere has helped to avoid the absorption of infrared energy from the Sun, so the planet's temperature has remained cool and stable relative to that of Venus. The Earth's atmospheric pressure is, by definition, 1 atmosphere at sea level, and it contains 78 percent nitrogen and 21 percent oxygen. There are small amounts of other constituents, including varying amounts of water vapor.

The diameter of Mars is 6,788 kilometers, which is considerably less than that of Earth or Venus. The density of Mars (3.92 grams per cubic centimeter) is considerably less than that of the other terrestrial planets. The core of Mars still consists mostly of iron, but it appears to have fewer dense elements such as sulfur. Mars has no magnetic field, so the core is believed to be solid. Older surface rocks in the southern hemisphere of Mars are magnetized, suggesting that the core was liquid early in the planetary history. Some of the surface of Mars, especially in the southern hemisphere, contains impact craters. Huge volcanoes have formed in places, mostly from the buildup of lava. There are large plains that are composed of lava flows, sedimentary rocks, and material ejected from volcanoes through the air. The surface also displays large fractures in the rocks in places. There is no evidence, however, that plate tectonics occur on Mars as on Earth. In some locations, sinuous valleys suggest that running water may have existed on the surface in the distant past. If water did exist on Mars, some form of life might also have existed. However, no evidence has yet been found that life currently exists on Mars. The main obstacles to the current existence of life on Mars arise from its low atmospheric pressure (only 1 percent that of Earth) and an atmsophere composed mostly of carbon dioxide. Also, air temperatures are cold, ranging from only 190 to 240 kelvins from night to day. The low atmospheric pressure of Mars means that no liquid water can currently be stable on the Martian surface; any water in solid form would quickly sublimate (change from solid directly to gas). Thus, the sinuous valleys could have been produced in the ancient past only when the atmospheric pressure was much higher.

KNOWLEDGE GAINED

Much of the information about Mercury, Venus, and Mars comes from observations from spacecraft that have flown by or landed on these planets. Also, some observations from Earth have been important.

Much of the information about Mercury came from the Mariner 10 spacecraft, which made three flybys in 1974 and 1975. Mariner 10 obtained information about the lack of an atmosphere, absence of a magnetic field, images of the abundant impact craters, and great variations in temperature.

Venus, with its continuous and thick cloud cover, has been difficult to study from Earth-based observatories. Mariner 2 flew by Venus in 1962, and the surface temperature and the lack of a magnetic field were detected. The Soviet Union's Venera missions determined that the atmosphere was mostly carbon dioxide. Venera 7 and 8 landed on the surface of Venus in 1970 and 1972, respectively. They measured the high temperature and pressure of the atmosphere and analyzed some of the surface rocks. Because of these conditions, however, the Venera spacecraft ceased to function within a few hours of landing. Later Venera spacecraft sent back color images of the Venusian surface, obtained better chemical analyses of the rocks, and made some high-resolution radar images of the surface. The Magellan spacecraft from the United States obtained almost complete high-resolution radar images of the surface, and they collected data on the gravity field.

Earth and its atmosphere have been studied in great detail by geologists, geophysicists, and atmospheric scientists for many centuries, so vast amounts of information have been collected compared to the other terrestrial planets. Geologists have mapped Earth's surface geology in great detail, and they have determined the mineralogy and chemical composition of the rocks at and near the surface of the Earth. Geophysicists have used seismic waves from earthquakes and data on variations in gravity and magnetism to estimate the composition of the interior of the Earth. Atmospheric scientists have determined the composition and variations of the present and ancient atmosphere of the Earth.

Initially Mars was viewed by telescope from the Earth, and some astronomers believed that they saw "canals" (or, from the Italian, "channels") on the surface. As a result, some nineteenth and early twentieth century astronomers speculated that Mars could have life. When Mariner 4 flew near Mars in 1965, it revealed that the southern hemisphere of Mars had impact craters and no canals. It also found that Mars had no magnetic field. Mariners 6 and 7 also flew by Mars in 1969, and they obtained more surface images. Mariner 9 orbited Mars in 1971. Detailed images of the surface revealed such features as large volcanoes and sinuous canyons. Vikings 1 and 2 landed on Mars in 1976, and they sent back information about the soil and air and searched for life until 1982, when they ceased to function. Numerous spacecraft since 1999 have extended the examination of Mars's surface and atmosphere. Chief among these have been the Mars Climate Orbiter, Mars Polar Lander, Mars Odyssey, Mars Pathfinder, Mars Exploration Rovers, Mars Express, Mars Reconnaissance Orbiter, and Mars Phoenix.

CONTEXT

There are still many questions about the terrestrial planets that need to be answered which should at least partially be answered by future space missions. For instance, the MESSENGER mission, launched on August 3, 2004, reached Mercury in 2008. It is designed to obtain further information about Mercury after it eventually enters orbit about the planet. The spacecraft needs to slow down considerably to be able to orbit Mercury, so the National Aeronautics and Space Administration (NASA) had it fly by Venus, Mercury, and Earth several times to slow it sufficiently to facilitate orbit of Mercury in March, 2011. As a result, MESSENGER flew by Venus in June, 2007, and took pictures to calibrate its cameras. It flew by Mercury in January 14, 2008, and again on October 6, taking a number of pictures of the surface. When MESSENGER eventually orbits Mercury, it will take more pictures of the surface to help determine the planet's geologic history, study the weak magnetic field to help determine its origin and the nature of the core, and conduct experiments to determine the composition of some of the rocks and atmosphere, which in turn will lead to a better understanding of their formation.

Another mission to Mecury, a joint effort between the European Space Agency and the Japanese Aerospace Exploration Agency, is the Bepi Colombo mission, to be launched in 2013 and expected to arrive in 2019. Bepi Colombo is expected to include not only an orbiter but also a lander, which would allow more detailed observations of the surface of Mercury.

Mars is undergoing continual intense investigation through both orbiting spacecraft and surface experiments conducted by rovers. In early 2009, Earth-based observations led to the discovery of a methane signature in the Martian atmosphere. The planned Mars Science Laboratory mission, expected to launch in 2011, may be repurposed to focus on this intriguing discovery.

Robert L. Cullers

FURTHER READING

Faure, Gunter, and Teresa M. Mensing. *Introduction to Planetary Science: The Geological Perspective*. New York: Springer, 2007. The authors have summarized the information about all bodies in the solar system in this textbook. Includes many photographs, graphs, and tables, as well as a glossary and an index. Best for those with some background in geology.

Hansen, V. L., and D. A. Young. "Venus's Evolution: A Synthesis." In *Convergent Margin Terranes and Associated Regions*, edited by W. Carlson et al. Boulder, Colo.: Geological Society of America, 2007. This paper summarizes much of what is known about Venus, including the formation of the impact craters and theories about the planet's formation. Contains figures to illustrate the discussion.

Marvin, Ursula. "Geology: From an Earth to a Planetary Science in the Twentieth Century." In *The Earth Inside and Out: Some Major Contributions to Geology in the Twentieth Century*, edited by D. R. Oldroyd. London: Geological Society, 2002. This article describes the geology of the solar system, starting with that of Earth, and how the understanding of other planetary bodies has elucidated our understanding of how Earth formed. Includes fourteen figures and several tables.

Selley, Richard, Robin Cocks, and Ian Plimer, eds. *Encyclopedia of Geology*. 5 vols. Oxford, England: Elsevier Academic Press, 2005. Contains a vast amount of information about geology, including articles about the planetary bodies of the solar system.

Solomon, Sean. "Mercury: The Enigmatic Innermost Planet." *Earth and Planetary Science Letters* 216 (2003): 441-455. Summarizes what is known about Mercury and what needs to be studied further. Outlines planned space missions. Provides many figures and other illustrations, including photos of the surface.

See also: Europa; Jovian Planets; Neptune's Interior; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Solar System: Origins.

Thermonuclear Reactions in Stars

Category: The Stellar Context

Thermonuclear reactions are the way stars generate energy for most of their lives. The various reactions are related to the changes stars undergo. They explain how the Sun has continued to shine for the past 4.5 billion years, and they predict how the Sun will change in the future.

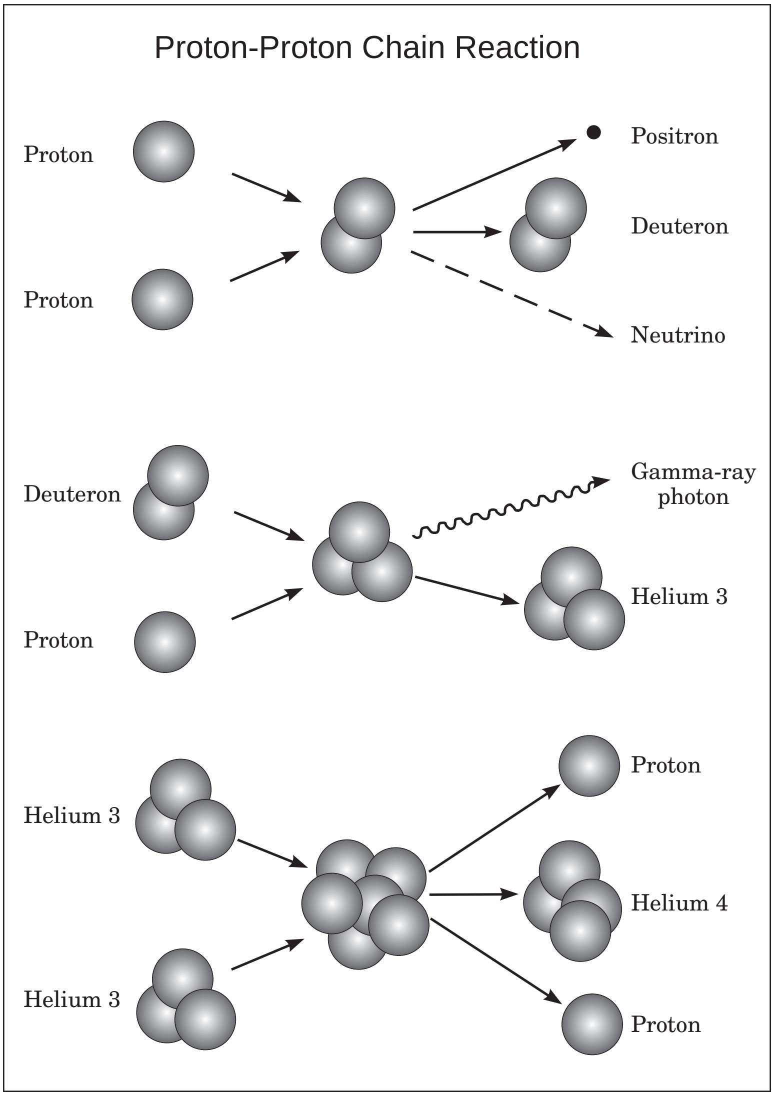
OVERVIEW

Our Sun is a main sequence star, about halfway through its lifetime (an estimated 10 billion years); as such, it is the closest stellar "laboratory" for research into the astrophysics and astrochemistry believed to take place in similar stars throughout the universe. The nature of the Sun's energy source baffled scientists for centuries. In 1905, however, while Albert Einstein was developing his special theory of relativity, an equation unexpectedly emerged: $E = mc^2$, where E is energy, m is mass, and c is the speed of light. This equation indicated that energy and matter are equivalent and can be transformed into each other. Another implication is that a small amount of mass is equivalent to a large amount of energy; for example, a 1-kilogram mass, if converted totally into energy, would yield 9×10^{16} joules. A joule is the energy needed to lift that 1-kilogram mass about 10 centimeters. The idea that mass could be converted into energy became the basis for thermonuclear weapons and for uranium-fueled fission reactors. It also is the mechanism that produces the energy of the Sun and other stars for most of their lives.

The Sun's power output, the rate at which it produces energy, is 3.83×10^{26} joules per second. One second of the Sun's energy output is equivalent to 10 million times the annual electricity consumed within the United States. To radiate at this rate, the Sun converts 4.25 billion kilograms of matter into energy every second.

The solar system, with its Sun, planets, satellites, and associated objects, developed from a cloud of gas and dust several light-years in diameter. That slowly spinning cloud started to contract under its gravitational attraction. As it shrank, the cloud rotated faster and spun off an equatorial disk. At the center, a ball of gas called the proto-Sun continued to contract and grow hotter. It was not yet a full-fledged star, since it was not converting mass into energy. Its rising temperature was caused by gravitational contraction of the gas and conversion of its gravitational potential energy into thermal energy. As the contraction continued, temperature, density, and pressure in the interior of the proto-Sun rose. When these became high enough, nuclear fusion ignited, and the proto-Sun became a star.

Since one cannot physically measure conditions within the interior of a star, astronomers mathematically model the Sun's internal structure. Using known conditions at the surface and the Sun's total mass, diameter, and energy output, computers can be used to calculate the changing conditions from the surface into the



Reactions start with the collision of two high-speed protons. Normally, as two protons approach each other, their positive electrical charge produces a mutual repulsion, and they slow down, stop, and then move away in a nearly elastic collision. If they are moving sufficiently fast, however, they can come close enough for the short-range strong nuclear force to allow the protons to attract each other. At this point, one of the protons is converted into a neutron, and a deuterium nucleus (hydrogen 2, an isotope of hydrogen with one proton and one neutron) results. Since electrical charge must be conserved, the change of a proton into a neutron results in the emission of a positron, a positively charged electron, which is the antimatter form of the electron. When a positron and an ordinary electron collide, they mutually annihilate each other, and their mass is completely converted into 1.6×10^{-13} joules, as described by Einstein's mass-energy equation. The energy is released as gamma rays, a high-frequency, high-energy form of electromagnetic radiation. A neutrino, a neutral particle with negligible mass, is also emitted by the nuclear fusion reaction to conserve energy and momentum simultaneously.

Next, a fast-moving proton collides with the deuterium nucleus, yielding a helium-3 nucleus and a gamma ray. The helium-3 nucleus has two protons and one neutron. Finally, two helium-3 nuclei collide to produce one helium-4 nucleus (with two protons and two neutrons) and two free protons. The net result of the interactions is that four protons fuse to yield one helium-4 nucleus, two positrons (which are annihilated when they encounter two free electrons), two neutrinos, and gamma rays.

interior. The model's values for mass and energy output must match those of the Sun for the model to hold validity. Calculations suggest that the core temperature is 15 million kelvins, pressure is as high as 200 billion atmospheres, and the density reaches 150 grams per cubic centimeter. It is under these conditions of high temperature, high pressure, and high particle density that the fusion of four protons (each a nucleus of ordinary hydrogen) into helium 4 occurs. It is a series of thermonuclear reactions, known as the proton-proton chain, in the 350,000-kilometer-diameter core that powers the Sun.

The mass of four protons is 6.6942×10^{-27} kilograms, while the mass of a helium-4 nucleus is 6.6466×10^{-27} kilograms. That mass difference of 0.0476×10^{-27} kilograms is converted into 4.28×10^{-12} joules of energy according to Einstein's mass-energy equation. The amount of matter converted per second by the Sun into energy is equivalent to 4 billion kilograms, or the mass of about 2 million automobiles. Each second, 8.9×10^{37} nuclear reactions transform 610 billion kilograms of hydrogen into 606 billion kilograms of helium and 3.8×10^{26} joules of energy.

The gamma-ray photons resulting from the reactions are absorbed and reradiated many times during their journey to the Sun's photosphere. This energy requires from 100,000 to 1 million years to make the 700,000-kilometer trip from the core of the Sun to its "surface," where most of the energy departs as photons of ultraviolet, visible, and infrared electromagnetic radiation.

The released radiative energy also prevents the Sun from collapsing. The gravitational attraction of the Sun's mass produces an inward pull, but the radiative energy of the Sun heats the interior gases. This produces an outward gas pressure to counter the inward pull of gravity. Fortunately for life on Earth, this has resulted in a star that has been stable for the past 4.5 billion years. With the remaining hydrogen in the core, our Sun should last for about another 5 billion years. At the end of that time, the core's hydrogen supply will be depleted, and the Sun's helium "ash" core will start to contract.

This core contraction stage will raise the internal temperature to the point at which the hydrogen in a shell around the contracting helium core becomes hot enough to fuse into helium 4. The Sun's outer layers will expand because of this increased energy production, and the Sun will enter a red giant stage in which it will engulf the inner planets of the solar system at least out to Venus and perhaps as far out as Mars. Later, as the core temperature increases to more than 100 million kelvins, two of the core's helium-4 nuclei will fuse into a beryllium-8 nucleus. Another helium-4 nucleus will fuse with the beryllium to produce carbon 12. In each case, a gamma-ray photon is emitted. Core reactions will cease as the helium 4 is depleted, but hydrogen and helium fusion in shells around the carbon core will continue. This ultimately will lead to expulsion of the Sun's outer layers to form a planetary nebula.

The Sun's carbon core will shrink to become a white dwarf about the size of the Earth. The Sun as a white dwarf will radiate only by its residual heat, since no thermonuclear reactions will occur within it. It will slowly cool to its final black dwarf phase, a burnt-out star with less than the original mass of the Sun.

Stars that form with masses greater than about 1.4 solar masses also fuse hydrogen into helium in their cores, but with a series of thermonuclear reactions known as the carbon-nitrogen-oxygen (CNO) cycle, rather than the proton-proton chain, due to core temperature, pressure, and particle density higher than those of a 1-solar-mass star. These conditions are caused by the larger mass of the star producing a greater gravitational pull inward. Because of higher temperatures and densities, the reactions will proceed at a faster rate. The star has a larger mass but consumes itself more rapidly. This results in a shorter lifetime for such a star. As the core depletes its hydrogen, it contracts and heats, and helium fusion commences. Hydrogen-to-helium fusion is initiated in the hydrogen shell around the core.

When the core's helium is expended, the core contracts again, with a resulting increase in temperature and density, and a new fusion reaction is ignited. This cycle of fusion, depletion, contraction, and new fusion continues until the nuclei in the core are ultimately converted into iron nuclei and the core is surrounded by several concentric shells of different fusing nuclei. The outer shell consists of hydrogen fusing to helium. Iron is the core's end product, since further fusion requires input of additional energy rather than its exothermic release. With no source of thermal energy to prevent the further gravitational collapse of the star, it implodes. Material bounces off the core, and the star violently explodes into a core-collapse supernova.

Energy is so abundant during the supernova explosion that nuclei are fused into elements with atomic numbers higher than iron, elements even as heavy as uranium and thorium. The remnant of the star collapses to as little as

10 kilometers in diameter and becomes a neutron star, because at such high pressures the electrons and protons of the star combine into neutrons. If a supernova leaves a stellar remnant with more than 3 solar masses, an even more dramatic end state occurs. The collapse continues beyond the neutron-star stage into a black hole. Nothing can stop the gravitational attraction that collapses the star to a size from which not even light can escape.

APPLICATIONS

The study of thermonuclear reactions is applicable to humanity's most urgent need: energy. The standard of living is governed by the availability of easily obtained energy, and humanity's present control of energy permits people to perform feats that previously were thought impossible. Human and animal muscle power, wind and water, coal, natural gas, petroleum, nuclear fission: Each has proved to be inadequate in one way or another, either because it is insufficient in amount or intensity, or because it produces harmful waste. Humanity needs a reliable source of energy, and thermonuclear reactions, either directly or indirectly, may well be that source.

Solar energy is the product of a natural thermonuclear reactor 150 million kilometers from Earth, and humans use solar energy in many forms without knowing it. In fact, except for tidal and geothermal energy, all energy sources are implicitly solar related. Coal, oil, and natural gas, for example, are forms of "fossilized" energy based originally on solar illumination of the planet. The direct application of solar energy is difficult, since solar energy is a diffuse energy resource. Large arrays of solar panels are needed to supply the required energy, and they operate only when the Sun shines.

One method of avoiding the earthbound problems of solar power is to construct in geostationary orbit around the Earth large solar power satellites. One such station, 5 kilometers by 15 kilometers in area, could supply the needs of New York City with plenty of energy to spare. Its solar cells would convert light into electricity, which would produce microwave radiation. That would be beamed to Earth, where a receiving antenna would convert the microwaves back into electrical power. One of these satellites would provide six to ten times the energy that an array of the same size on Earth would provide. Several hundred of these satellites would supply a large portion of humanity's energy needs.

Another possibility is to design machines that would permit the control of fusion reactions on Earth. For decades, engineers and scientists have struggled with the problem of obtaining sufficiently high particle densities, temperatures, and pressures. The main problem is the design of a vessel that will constrain deuterium long enough for the fusion to occur. Magnetic fields and lasers have been employed to initiate the fusion process, but the goal of a net energy output on an economic scale remains elusive. When practical fusion reactors are developed, hydrogen from the water of the Earth's oceans will provide an abundant source of fuel.

Scientists who construct models of the thermonuclear reactions and other processes that occur in the cores of stars use the Sun as their test case. Certain predictions about the Sun can be made from the model. One of these involves the number of neutrinos emitted by the reactions in the Sun's core. Neutrinos are particles with little mass and no electrical charge, traveling at speeds close to that of light. They do not interact readily with matter; a neutrino can easily pass through a light-year-thick layer of lead shielding. The model predicts the number of solar neutrinos that should be counted by experiments designed to detect them. The first measurements in the 1970's detected only one-third of the expected number of neutrinos. Scientists questioned if the models of the Sun's interior processes were incorrect, or if there was something unknown happening in the Sun's interior that the models had not taken into account. After decades of uncertainty over the Sun's missing neutrino flux, it has been determined that solar neutrinos change into other forms or flavors during their trip between the Sun and Earth, and the original apparent discrepancy seems to have been resolved.

Since neutrinos make the trip from the core to the detectors at nearly the speed of light, they provide information about what is occurring now in the Sun's core. The electromagnetic radiation from the Sun's photosphere discloses what occurred in the core 100,000 or more years ago. Further theoretical and experimental work should produce an answer on determining if the Sun's energy output varies substantially over time. Indeed, many scientists think of the Sun as a variable star, but fortunately that level of variability is quite low compared to other "traditional" classes of variable stars.

CONTEXT

In the eighteenth century, Immanuel Kant estimated that if the Sun were composed of coal, it would burn only for several thousand years. In the nineteenth century, Hermann von Helmholtz and Lord Kelvin independently reasoned that gravitational contraction of the Sun could provide its energy by converting gravitational potential energy into thermal energy. This theory allowed the Sun's age to be increased to 20 to 50 million years, but it did not satisfy geologists and the biologists, who argued that hundreds of millions to billions of years were necessary for the evolution of life and the geophysical development of the Earth.

In the nineteenth century, physicists conducted experiments to determine how the speed of light changed as the speed of the medium through which it traveled was varied. The conclusion was that there is no medium through which light moves. It was also observed that the speed of light is constant, no matter how fast its source moves. Classical physics was unable to accept this seemingly nonsensical result for the speed of light. Einstein, however, stated that the speed of light is a constant in any inertial frame of reference, and went on to investigate the consequences of this postulate. While making those derivations, his famous mass-energy equivalence emerged. In 1939, Hans Albrecht Bethe and Carl von Weizsäcker hypothesized that nuclear reactions could generate the Sun's energy. They suggested that four protons could fuse into helium 4, and that the mass difference was converted into energy.

In the early 1950's, the thermonuclear or hydrogen bomb was developed, in which ignition of a fission bomb trigger produces high temperatures that lead to fusion of deuterium and the release of even more energy. This produces the hydrogen bomb's greater destructive power and in a sense replicates in an uncontrolled fashion the tremendous power of the Sun's fusion process.

Stephen J. Shulik

FURTHER READING

Chaisson, Eric, and Steven McMillan. *Astronomy Today*. 6th ed. New York: Addision-Wesley, 2008. This astronomy textbook contains a thorough description of the stages of stellar evolution, complete with transparent overlays, and an exceptionally lucid treatment of thermonuclear energy generation.

Dinwiddie, Robert, et al. *Universe*. New York: DK Adult, 2005. A remarkable collection of articles written by science writers and professional astronomers on a wide range of topics that span the discipline of astronomy. Heavily illustrated and filled with high-quality photographs. For the general reader.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Provides a good description of stellar evolution and thermonuclear energy generation.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook. Offers a good description of stellar evolution and thermonuclear energy generation.

Hansen, Carl, Steven Kawaler, and Virginia Trimble. *Stellar Interiors: Physical Principles, Structure, and Evolution*. 2d ed. New York: Springer, 2004. Covers the fundamentals of stellar physics, structure, and life cycle. Newer research topics such as astroseismology and the effect of magnetic fields are also covered. For advanced undergraduates.

Prialnik, Dina. *An Introduction to the Theory of Stellar Structure and Evolution*. Cambridge, England: Cambridge University Press, 2000. An undergraduate textbook covering all aspects of stellar evolution and the structure of stars. Full solutions to exercises as well as basic physics and mathematics are included.

Schneider, Stephen E., and Thomas T. Arny.

Pathways to Astronomy. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses. Divided into lots of short sections on specific topics, it presents a thorough discussion of stellar evolution and thermonuclear energy generation.

Thornton, Stephen T., and Andrew Rex. *Modern Physics for Students and Engineers*. 3d ed. New York: Brooks/Cole, 2005. A comprehensive presentation of the development of relativity, quantum mechanics, and nuclear and particle theory and experimentation. For undergraduates and serious scientific researchers.

Young, Hugh D., and Roger A. Freedman. *University Physics with Modern Physics*. 11th ed. New York: Addison-Wesley, 2003. An undergraduate text that spans classical mechanics, thermodynamics, Maxwell's electrodynamics, optics, and modern physics. Offers a good introduction to aspects of elementary particle physics as well.

See also: Brown Dwarfs; Gamma-Ray Bursters; Main Sequence Stars; Novae, Bursters, and X-ray Sources; Nuclear Synthesis in Stars; Protostars; Pulsars; Red Dwarf Stars; Red Giant Stars; Stellar Evolution; Supernovae; White and Black Dwarfs.

Titan

Categories: Natural Planetary Satellites; The Saturnian System

Saturn's largest satellite, Titan, is the only satellite in the solar system with a thick atmosphere. Astronomical observations made from Earth established some time ago that Titan has a density slightly greater than compressed ice, indicating a composition primarily of ice but with a relatively small rocky core. Observations made by the Cassini-Huygens spacecraft show a surface with multiple hydrocarbon lakes that could be breeding grounds for primitive living organisms.

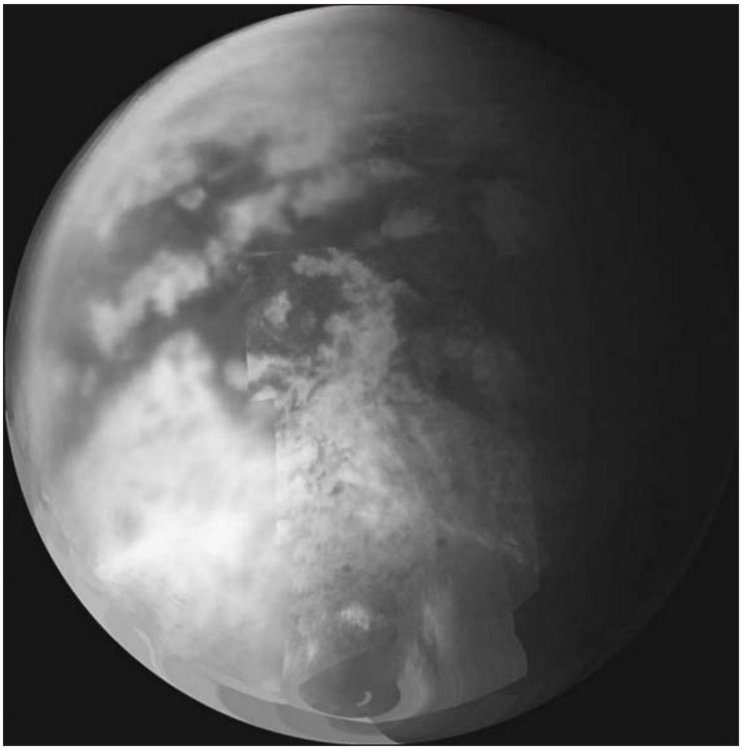
OVERVIEW

Titan is Saturn's largest satellite, with a diameter of about 5,150 kilometers. Its atmosphere (whose density is several times that of Earth's atmosphere) was discovered in 1944 from spectral analyses of sunlight reflected from the cloud cover. The spectral data indicated the presence of methane gas (CH_4). Additional Earth-based observations in 1973 showed a reddish, hazy atmosphere, which was assumed to be photochemical smog created by ultraviolet light from the Sun acting on the methane and other hydrocarbon compounds.

Because Titan is such an unusual satellite, the Pioneer 11 spacecraft flew by Titan in 1979, followed by Voyager 1 and 2 in 1980 and 1981, respectively. Unfortunately, their instruments were not sensitive enough to penetrate Titan's thick atmosphere, although it was learned that its major constituent is nitrogen, with methane and smog making up less than 10 percent of the atmosphere. Hazy smog is formed as the methane is catalyzed by ultraviolet light from the Sun to form more complex organic molecules, similar to the manner in which photochemical smog is produced from unburned fuel in the exhaust emitted by vehicles in Earth's large cities. Because Titan is quite far from the Sun and relatively little of the available light would penetrate the thick clouds, the surface temperature was predicted to be about 94 kelvins.

Calculations indicated that at these temperatures methane should condense from the clouds and fall as rain. The denser organic molecules created in the atmosphere, such as ethane (C_2H_6), would also eventually settle on the surface as a layer of malodorous slime. It was thus assumed that the icy surface was covered by either an ocean of liquid methane or a hydrocarbon swamp, with frequent rainstorms of methane.

In 1994, scientists used the Hubble Space Telescope at near-infrared wavelengths (where the haze is more transparent) to map some of Titan's surface features according to their reflectivity. Although details were not resolvable, light and dark surface features were recorded over Titan's sixteen-day rotation period, and one bright area the size of Australia was documented. Definitive conclusions about the na-



A composite image of Titan from the Cassini spacecraft. (NASA/JPL/University of Arizona)

ture of the dark and bright areas could not be ascertained, but images proved that the surface was not a global ocean of methane and ethane, as had been assumed; at least part of the surface is solid. Although definite conclusions could not be made, it was thought that the bright areas were major impact craters in the frozen surface. Information gleaned from this research provided important background information for the Cassini mission, the program of National Aeronautics and Space Administration (NASA) that sent a robotic spacecraft to study Saturn and its satellites. In addition to data to be gathered from flybys, Cassini would release the European Space Agency's Huygens probe, which would parachute to the surface. Images of Titan from the Hubble telescope were used to locate an optimum landing site and to predict how Titan's winds would affect the parachute as it descended through the atmosphere.

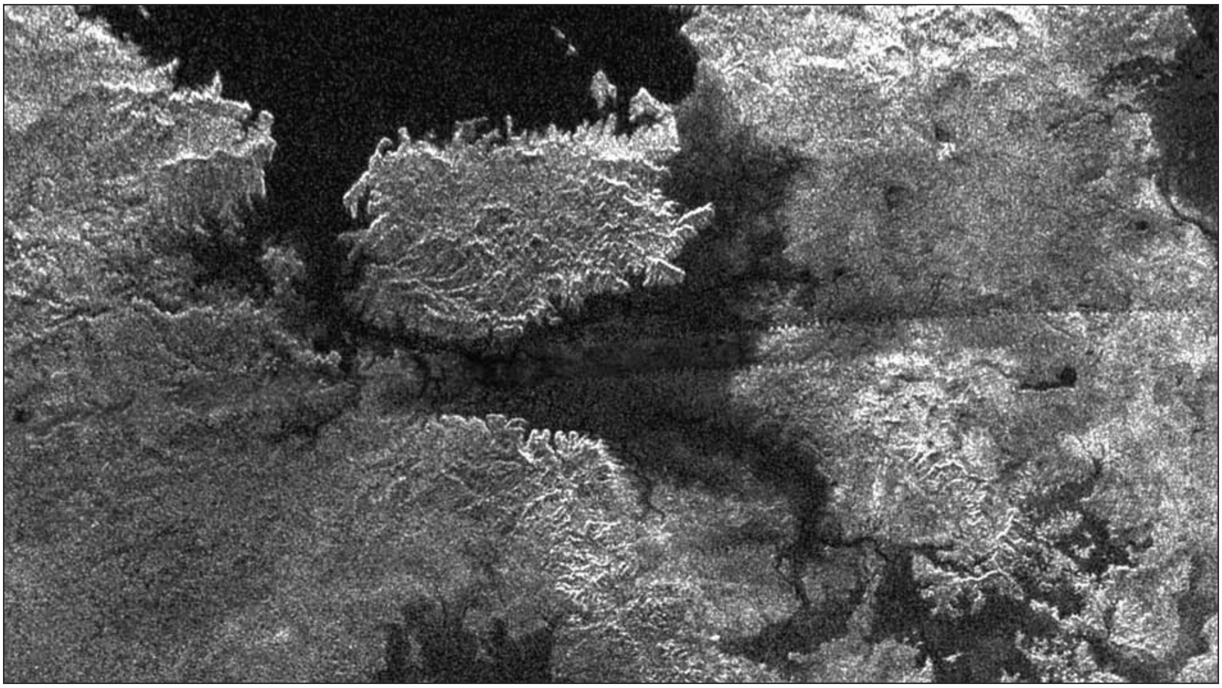
The Cassini spacecraft was launched in October, 1997, for its seven-year voyage to rendezvous with Saturn. Beginning to orbit Titan in 2004, it flew 1,192 kilometers above the surface, using infrared cameras and radar to produce detailed maps. It detected irregular highlands and smoother dark areas, including one large region, about the size of Lake Ontario (232 by 72 kilometers), so reminiscent of a lake that its perimeter even exhibited sinuous drainage channels leading to an apparent shorelike boundary. Because the surface temperature was so cold (94 kelvins), the lakes were presumed to be liquid methane and ethane fed by streams of dark organic gunk washed by precipitation from the highlands. Methane evaporating from the lakes would replenish the methane in the atmosphere, from which it would eventually precipitate and return to the surface as rain, mimicking the hydrologic cycle on Earth. The fact that this feature appears in Titan's cloudiest region, where presumably storms are intense enough that methane rain reaches the surface, gave credence to the lake hypothesis. Furthermore, Titan's cold temperature would require a long time for liquid methane on the surface to evaporate; thus, a methane-filled lake would remain stable for a long time.

The Huygens probe was released on January 14, 2005. As it descended, it recorded the temperature, pressure, wind speed, and atmospheric composition at regular time intervals. It also radioed back more detailed images of the surface, showing dark drainage networks leading into the smooth areas but relatively few craters, as expected. The bright spots appeared to be "islands" around which dark material had flowed in the past. Other images showed areas evocative of water ice extruded onto the surface and short, stubby, dark channels which could be springs of liquid methane. Although these data suggested running liquids, no clear evidence of liquid methane was detected on the surface. After landing, Huygens probed the surface, which had the consistency of wet sand covered with a thin crust, possibly consisting of ice mixed with small amounts of solid methane. First pictures of the surface showed a plethora of small erosion-rounded pebbles, assumed initially to be rocks or granite-hard ice blocks, on an orange-colored surface. They later were determined to be mixtures of water and hydrocarbon ice. One image pictured tendrils of surface fog, presumed to be ethane or methane.

The Huygens lander took this image on January 14, 2005, from Titan's surface. The rocks at the front are about 4 to 6 inches in diameter and may be the remains of a lakebed. (ESA/NASA/JPL/University of Arizona)

The concentration of methane in Titan's atmosphere is puzzling, because ultraviolet light from the Sun would dissociate the methane into carbon and hydrogen, which would either react with the nitrogen to form ammonia (NH_3) or be dissipated into space. More complex organic molecules, such as ethane, would also be created; being heavier, they would settle to the surface. It has been calculated that atmospheric methane should remain in the atmosphere for fewer than one million years. Consequently, methane must be injected from some surface source. Perhaps it is outgassed from the methane in the icy crust. Another possibility is a methane volcano. Detailed observations have identified one area where ice and methane may be rising to the crust from a subterranean heat source, to form a methane volcanic caldera emitting methane gas. Since Titan is too small to have a molten interior, the heat source driving the release of methane gas is suspected to be tidal heating, the frictional force generated as this massive satellite revolves in its elliptical orbit about Saturn. Several dark surface markings having straight boundaries with preferred orientations suggest the presence of internal tectonic processes.

Analyses of close flybys of Titan made by the Cassini orbiter in 2008 as well as conclusions based on data gathered by flybys made between 2004 and 2006 provided suggestive evidence for the possibility that Titan's surface may have active cryovolcanoes, perhaps spewing water, ammonia, and methane. Cassini recorded variations in brightness and reflectance (the ratio of reflected light to the incident light upon a surface) in two separate regions of Saturn's largest satellite using its Visible and Infrared Mapping Spectrometer. In one region, the reflectance increased significantly and remained at the elevated level; in the other, it rapidly increased and then tailed off again. Both would indicate vapor or liquid being ejected out of an active vent. Such a mechanism would explain why Titan continues to maintain a thick methane atmosphere when, without replenishment, it should have been greatly diminished over the passage of geologic time.



Cassini captured this image of Titan's surface on February 22, 2007, with a resolution of about 700 meters. The view looks directly down on an island (possibly a peninsula), about the size of the Big Island of Hawaii, in the middle of one of the moon's hydrocarbon lakes. (NASA/JPL)

KNOWLEDGE GAINED

Using radar, Cassini was still mapping Titan's surface in 2009 and was expected to continue doing so for some time to come. The hundreds of observed dark areas are believed to be lakes of liquid methane or ethane more than 12.2 meters deep, while shadowy dunes running along the equator are assumed to consist of complex solid organics. Titan's surface seems to contain many gigantic organic chemical factories producing complex hydrocarbons in an abundance surpassing all of Earth's oil reserves. The amount of liquid on Titan's surface is important to ascertain, because methane is a strong greenhouse gas; without atmospheric methane, Titan's surface would be even colder. Liquid methane on the surface could remain, at most, a million years before dissociating and reacting to form heavier hydrocarbon compounds. It is believed that the atmospheric methane is constantly being supplied by volcanic eruptions from the mantle.

In late December, 2008, after much analysis, a group of researchers were ready to publish a scientific article in the research journal *Icarus* reporting the first image taken of a liquid on a planetary surface other than the Earth's. The Mars Phoenix lander had provided clear evidence about six weeks earlier of subsurface water ice at its far north landing site, but when that ice was exposed, it fairly quickly sublimated into the gas phase and became part of the Martian atmosphere. However, an image taken by the Huygens probe after reaching the surface of Saturn's satellite Titan appeared to have clearly recorded a droplet of methane near the edge of the robotic spacecraft itself. The small droplet might have been created by heat emanating from the probe when it condensed humid air to temporarily form liquid methane. In several other images, splotches that appeared and then were not seen in subsequent images of the same area were believed by the authors of the *Icarus* paper also to be droplets of methane.

Titan has an extensive atmosphere, including a methane layer extending 696 kilometers above the surface. There the methane molecules are dissociated by ultraviolet light to form ethane (C_2H_6) and acetylene (C_2H_2). Cassini images showed two thin haze layers. The outer haze layer, floating about 400 kilometers above the surface, is where additional molecules (such as hydrogen cyanide) are formed from carbon, hydrogen, and nitrogen. About 200 kilometers above the surface, there is a thick global smog of complex organic molecules, produced by chemical reactions among the hydrocarbons dissociated by ultraviolet light. This haze layer absorbs about 90 percent of the incident sunlight, leaving only an orangish haze to reach the surface. It is not currently understood why two separate haze layers are present. Although Titan has a dense atmosphere, it is relatively inefficient at reradiating infrared radiation, thus producing negligible greenhouse warming.

CONTEXT

The Cassini-Huygens mission was a joint venture of NASA, the European Space Agency, and the Italian Space Agency. Enough data were gleaned to keep researchers occupied for years to come.

Titan's surface temperature (94 kelvins) appears to make the satellite a place inhospitable for life to evolve. This environment, although colder, is remarkably similar to that found on Earth billions of years ago, before life began adding oxygen to the atmosphere. Four billion years ago, Earth was covered with warm, shallow seas containing hydrogen, ammonia, and methane gases. From this primordial soup, driven by ultraviolet light and lightning discharges, complex hydrocarbons, including amino acids, formed. Over time, the amino acids linked together to form proteins, eventually creating one that was able to replicate itself. At that point, life was created and molecular evolution became biological evolution.

If life could evolve in Earth's primordial soup, it seems reasonable to suppose that the pools of organic gunk on Titan's surface could form amino acids, if not self-replicating proteins. Studying Titan's prebiotic chemistry can therefore facilitate the understanding of how life may have originated in the universe.

George R. Plitnik

An artist created this image of a lake and smoggy atmosphere of Titan based on data that led scientists to conclude the moon has lakes of liquid hydrocarbons. (NASA/JPL)

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. This easily accessible work, written for laymen with inquisitive minds, has an excellent summary of the latest knowledge about Titan as well as pictures from the Huygens landing and an instructive graph of the variation of pressure and temperature as a function of altitude.

Coustenis, Athena, and Fredric W. Taylor. *Titan: Exploring an Earthlike World*. Hackensack, N.J.: World Scientific, 2007. A revised and expanded edition of the 1999 title *Titan: The Earthlike Moon*, this volume summarizes all that is known about Titan through the Cassini-Huygens mission, by two of the project's investigators. Aimed at a general audience, but scientifically rigourous nonetheless.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. This authoritative and regularly updated text considers all the major planetary objects in our solar system. The material is presented by grouping objects under unifying principles, thus elucidating their similarities and their differences as well as the physical processes behind their evolution. Although most of the material is descriptive, some algebra and elementary calculus are included.

Lorenz, Ralph, and Jacqueline Mitton. *Titan Unveiled: Saturn's Mysterious Moon Explored*. Princeton, N.J.: Princeton University Press, 2008. This illustrated tome was the definitive work covering everything known about the surface and atmosphere of Titan at the time of publication. Because Cassini-Huygens was still years away from its encounter with Titan, the authors had to predict—but with some accuracy it turned out—features and conditions on the surface based on limited data.

Sagan, Carl. *Cosmos*. New York: Random House, 1980. Based on the television series of the same name, this lavishly illustrated book includes not only information about Titan's atmosphere but also speculation about the possibility of a methane-based life-form evolving in this environment.

Seeds, Michael A. *Foundations of Astronomy*. 9th ed. Belmont, Calif.: Thomson Brooks/Cole, 2007. This well-illustrated text commingles experimental evidence and theory to provide deep, but well-explained, elucidations of many fascinating facets of the universe. Although only two pages are devoted to Titan, there are four pictures, one of which was taken from the surface by the Huygens probe.

See also: Enceladus; Iapetus; Planetary Ring Systems; Planetary Satellites; Saturn's Magnetic Field; Saturn's Ring System; Saturn's Satellites; Solar System: Element Distribution.

Triton

Categories: Natural Planetary Satellites; The Neptunian System

Triton, Neptune's largest satellite, is the solar system's only major satellite that is in a retrograde orbit. It has smoke plumes that astronomers and planetologists cannot explain. It appears to be younger than most of the satellites and planets in the solar system.

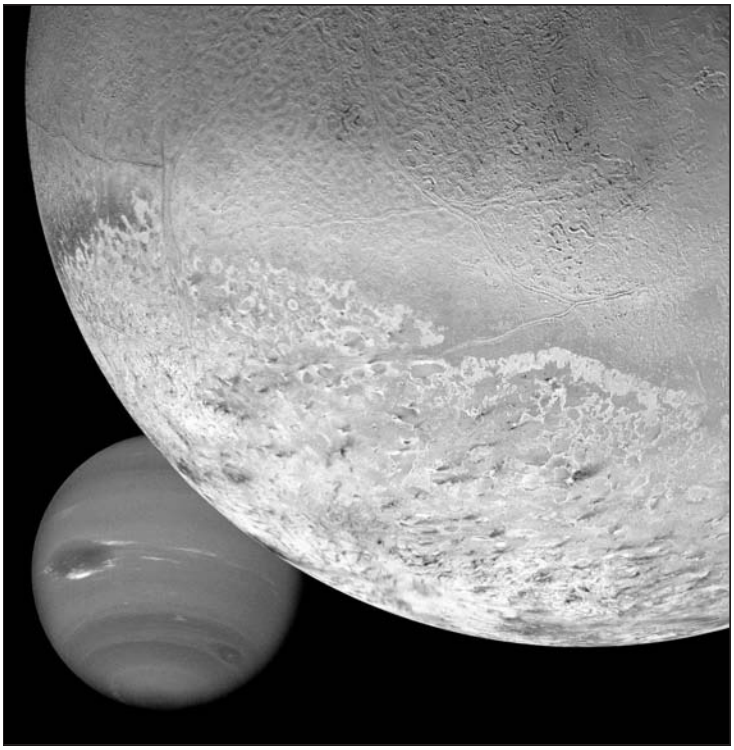
OVERVIEW

The Neptunian satellite Triton is 2,706 kilometers in diameter. The satellite's density is 2.07 grams/centimeter^3. With this density, models of the satellite can be constructed. It is thought that there is a metallic core, a silicate mantle, a layer of ice, a possible ocean, and a top layer of ice. The core is expected to have a radius of about 600 kilometers and the mantle a thickness of about 350 kilometers, with a 150-kilometer layer of ice below the ocean and a 250-kilometer layer above the ocean.

William Lassell, a brewer by trade, found Triton on October 10, 1846, using a telescope he built himself. An amateur astronomer, Lassell had been asked by Sir John Herschel to look for satellites of the newly discovered planet, Neptune. Triton is 355,000 kilometers from Neptune. The eccentricity of Triton's orbit is zero, meaning that the orbit is circular. Triton is a most unusual satellite. Its orbit is retrograde, that is, it rotates around Neptune in the opposite direction to the rotation of Neptune. It is synchronous with Neptune, meaning it presents the same face to Neptune at all times. Being synchronous also means that the rotation time of Triton is the same as the time for one orbit around Neptune, 5 days and 21 hours. The angle of inclination is 23°. The angle between the orbit of Triton and the equator of Neptune is 23°.

This large angle of inclination, coupled with the retrograde direction of Triton's orbit, suggests that Triton was captured by Neptune's gravitational field. The method by which this capture occurred is unknown. Triton may have collided with another satellite, causing Triton to slow down enough to be captured and destroying the other satellite at the same time. Another idea is that the other satellite was knocked out of orbit, and Triton was then captured. A third idea is that Triton was part of a binary system. Triton was captured; however, the partner escaped. When Triton was captured, the orbit was probably highly elliptical. The gravitational force of Neptune gradually changed Triton's orbit into the current circular orbit. During the period of this change, the strong pull when Triton was close and the weaker pull as Triton was far away would have caused tidal flexing, or movement within Triton's structure. Such motion would have caused internal friction, generating heat. The heat would have caused differentiation, that is, separation of the components of Triton. Heavier materials, such as metals, would have sunk to the core; medium-mass materials, such as silicates, would have formed a mantle; and lighter materials would have been forced to the surface.

Many believe that Triton is a volcanic satellite because of plumes of dark material that appear to be blown from its surface into the air. This material is concentrated enough to be easily seen. The plumes are about 2 kilometers across, rise as high as 8 kilometers, and consist of particles that probably are less than two millimeters in diameter. The size of the particles can be inferred because they do not settle to the surface. These plumes may be putting out 10 kilograms of material per second and may last for years.



Triton (foreground) and Neptune appear together in this montage of images from Voyager 2. (NASA/JPL)

Triton has a thin atmosphere, which is composed predominantly of nitrogen at a pressure of 14 millibars (Earth's atmospheric pressure is about 1 bar). There are clouds composed of condensed nitrogen. A diffuse haze can also be seen. The haze probably consists of hydrocarbons and nitriles, produced by the action of sunlight on methane. Wind-driven streaks are oriented in an east-west direction. The wind causes some of the streaks by material blown from the plumes. When all of the streaks, clouds, and plumes are considered, the winds blow northeast close to the surface, eastward at intermediate levels, and westward at the top of the troposphere.

Voyager 2 was 530,000 kilometers from Triton when it took this image in 1989 through green, violet, and ultraviolet filters. (NASA)

Infrared technology gave scientists the first look at Triton's surface. The spectra could be modeled only by a combination of solid methane (CH_4), called methane ice; liquid nitrogen; and water ice. Later spectra showed nitrogen (N_2) in a solid or liquid form. One idea is that there is a sea of nitrogen with small amounts of dissolved methane. More likely, there is a layer of solid nitrogen with contaminants of methane, carbon dioxide (CO_2), and carbon monoxide (CO). Even at the measured temperature of 38 kelvins, the nitrogen ice will sublime, forming the thin atmosphere first found by Voyager 2. The nitrogen refreezes at the winter pole of Triton, causing a polar ice cap. Solid nitrogen is very transparent; therefore, the sunlight that does reach Triton can heat the interior of the solid nitrogen in a greenhouse effect. Nitrogen, in either gaseous or liquid phase as a result of heating nitrogen originally in solid form, will flow to the surface, where it then freezes. Some of the nitrogen will escape into the atmosphere. The layer of nitrogen thus moves, or at least thins, with the seasons. There is plenty of time for the seasonal shift, because Triton has a 688-year climate cycle due to its unique rotational and orbital motion. Thus, there is a higher albedo (0.7) on the winter end of the satellite, where the layer of nitrogen ice is thick, and a lower albedo (0.55) on the summer end of Triton. The summer end will show methane ice, which has a reddish color. Radiation will eventually turn the methane dark. The fact that the methane is not dark means that the methane ice is refreshed on a short timescale.

Triton displays at least three different types of surfaces: a bright polar area; areas of dark patches surrounded by lighter material; and high, walled plains. The bright polar area seems to be nitrogen ice on top of a cantaloupe-like, dimpled surface. The dimples, called cavi, are caused not by volcanoes but by diapirism. Diapirism is generated by a gravitational instability in which less dense material flows up through denser material. The density gradient may be caused by temperature or by a difference in composition. The implication on Triton is that the crust has distinct layers, and that the top layer is no more than 20 kilometers thick. The area also has linear ridges across the cantaloupe terrain.

Dark patches within the lighter material are called maculae and probably are composed of carbonaceous material, such as methane ice. Bright material is probably nitrogen ice. Maculae may mark spots of heat that have lost the nitrogen ice layer, allowing the methane ice, with its lower albedo, to show through.

The high plains are caused by a flow of volcanic ice. Some of the plains are smooth plains with a flat-to-undulating structure. Other plains are surrounded by a terraced wall or steppes, called scarps. These plains are very flat, implying that they were filled with liquid at one time. Scarps may be the remainder, as the material on the plains sublimed.

One unique feature of Triton is its small numbere of craters. There is one crater that is 27 kilometers across, named Mazomba, but the small number of craters suggests either that the surface of Triton is very young or that the surface must have been refreshed fairly recently—or both. Part of Triton's surface is considered cryovolcanic instead of silicate-magma volcanic. Cryovolcanic activity is the eruption from the subsurface of icy-cold liquids, which then refreeze on the surface in a more or less smooth structure.

KNOWLEDGE GAINED

Much of what has been discovered about Triton has come from Earth-based instruments and the Hubble Space Telescope (HST). Hale Observatory used narrow-band spectrophotometry to determine that Triton had a constant spectral reflectance. Astronomers have compared data from HST, the Infrared Telescope Facility at the University of Hawaii, and Voyager 2 over a period of years to see if there is a seasonal change in Triton's surface. It appears that there is a change, but since the climate cycle is so long, the data remain inconclusive. Both types of surface composition, methane ice and solid nitrogen, were detected by infrared spectra from an Earth-based instrument.

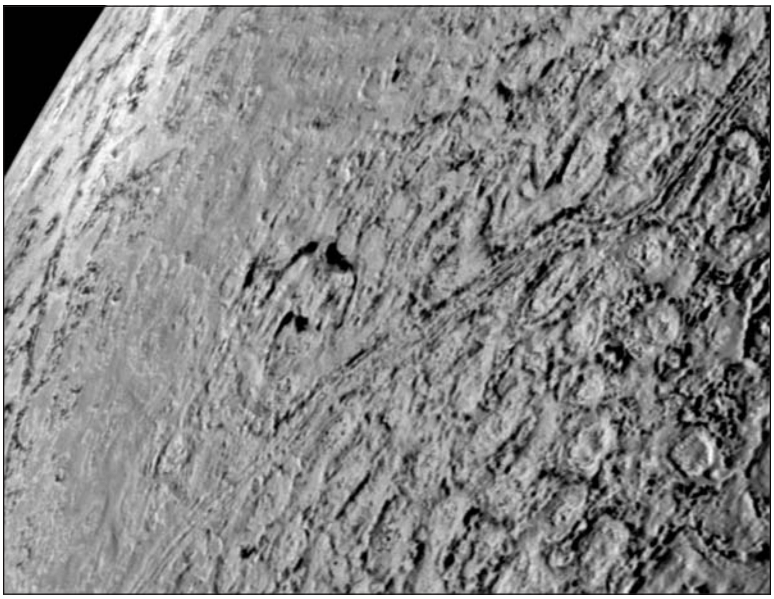
Voyager 2 provided more information in a short time than the land-based instruments had been able to gather in the years since Triton's discovery. The spacecraft's small changes in flight path caused by Triton's mass allowed that mass to be calculated. Pictures allowed the size to be determined. Density could then be calculated. Models of the structure of the satellite could then be developed. The density, 2.07 grams/centimeter^3, indicates that there is a large component of silicate materials, even though they do not show in infrared spectra because they are under ice. Voyager also detected the nitrogen atmosphere.

Pictures showed the effects of wind on Triton, a phenomenon that was unexpected. Varied terrain and plumes were also noticed in the pictures. The temperature measured was the coldest of any surface measured. Even with the cold temperature, the different terrains indicated that the surface had been refreshed more recently than any other moon or planet except those planets or moons that are geologically active.

CONTEXT

The density of Triton is close to that of the Pluto-Charon system. Is that a coincidental fact, or are Triton and Pluto related? Could Pluto at one time have been a satellite of Neptune that was knocked off by Triton? The orbital inclination, rotational speed, and retrograde motion all point to some cataclysmic occasion that produced Triton as a satellite.

Triton's surface features raise interesting questions about its energy source. Triton's is the coldest surface in the solar system, yet it also appears to be active, given the plumes and smoothness observed. How can these conditions coexist? The idea that enough sunlight penetrates a deep sheet of solid nitrogen to produce a greenhouse effect under the ice is startling but



Voyager 2's flyby of Triton produced this close-up image (from 40,000 kilometers), showing the moon's unique northern hemispheric terrain of more or less regularly spaced circular depressions ringed by ridges. These are not impact craters but more likely areas of collapsed ice. (NASA/JPL) may be true; it does appear that the ice sublimes and then refreezes in another place. Yet where does Triton get the energy to produce plumes rising 8 kilometers into the satellite's tenuous atmosphere? Even given the satellite's rather low gravitational acceleration, this phenomenon completes the satellite's overall mystery. Is Triton's interior heated radiogenically? Is there another heat source?

It is possible that scientists do not understand as much about the effect of very low atmospheric pressure, low temperature, and heavy mass as was previously thought, because the planet's heat is generated somewhere. The theory that Triton is heated radiogenically will have to wait until the subsurface can be monitored for radioactive isotopes, or for their daughter isotopes, before it can be confirmed or disproved. Certainly the answer to the question of Triton's energy source will add both to our understanding of the origins of the solar system and to our knowledge of energy physics.

C. Alton Hassell

FURTHER READING

Bond, Peter. *Distant Worlds: Milestones in Planetary Exploration*. New York: Copernicus Books, 2007. The author discusses each of the planetary systems, including planets, moons, and rings. Exploratory space missions and how they have developed our knowledge of each system are also addressed. Illustrations, bibliography, appendix, index.

Corfield, Richard. *Lives of the Planets*. New York: Basic Books, 2007. The author takes the reader through the different planets and the information gathered by space missions that investigated them. Index.

Croswell, Ken. *Ten Worlds: Everything That Orbits the Sun*. Honesdale, Pa.: Boyds Mills Press, 2007. Basic information on each system is presented separately. Illustrations, bibliography, index. For younger readers.

Cruikshank, Dale P., ed. *Neptune and Triton*. Tucson: University of Arizona Press, 1995. Voyager 2's 1989 encounter with Neptune revealed Triton to be a frozen, icy world with clouds, haze layers, and vertical plumes of particles rising high into the thin atmosphere. Originally presented as papers at a 1992 conference, the chapters in this volume are all by experts on Neptune, its many satellites, and its near-space environment. Until engineers can design propulsion systems for the next mission to the outer solar system, this 1,249-page tome will remain the most authoritative one-volume resource on Neptune and its satellites.

Dasch, Pat. *Icy Worlds of the Solar System*. Cambridge, England: Cambridge University Press, 2004. This book discusses ice, first on Earth, then on other solar-system bodies, including Triton. Illustrations, bibliography, index.

Hartmann, William K., and Ron Miller. *The Grand Tour: A Traveler's Guide to the Solar System*. 3d ed. New York: Workman, 2005. Focusing on the Voyager missions, this volume addresses each major planet and the major moons, including Triton. Includes outstanding illustrations. Illustrations, bibliography, index.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Accessible to the serious general audience.

Lopes, Rosaly M. C., and Michael W. Carroll. *Alien Volcanoes*. Baltimore: Johns Hopkins University Press, 2008. The focus is on volcanism throughout the solar system, including the possibilities on Triton. Illustrations, bibliography, index.

McFadden, Lucy-Ann Adams, Paul Robest Weissman, and T. V. Johnson, eds. *Encyclopedia of the Solar System*. San Diego: Academic Press, 2007. The editors have collected articles written by many experts. It is one of the best surveys of material about the solar system. Illustrations, appendix, index.

See also: Neptune's Atmosphere; Neptune's Great Dark Spots; Neptune's Interior; Neptune's Magnetic Field; Neptune's Ring System; Neptune's Satellites; Planetary Satellites; Uranus's Satellites.

U

Ultraviolet Astronomy

Category: Scientific Methods

Ultraviolet astronomy explores the universe by focusing on wavelengths of the electromagnetic spectrum that are shorter than those of visible light. This portion of the spectrum is particularly important to astronomy, as practically all stars and many of the most abundant elements in the universe emit energy in the ultraviolet range.

OVERVIEW

Any material with a temperature above absolute zero emits electromagnetic radiation, and that radiation carries with it information about the nature of the event that produced it. An object will emit radiation over a range of wavelengths, with a concentration at a single wavelength. Very hot objects produce shorter wavelengths, while cooler objects emit longer wavelengths. For example, as metal is heated, it first glows red (the longer wavelengths of visible light), then, as its temperature increases, it begins to glow in the shorter-wavelength yellow light. In space, objects that are very cold, perhaps only a few degrees above absolute zero, will emit radiation in the very long infrared and radio wavelengths. At the other extreme, very hot stars give off ultraviolet radiation, X rays, and gamma rays.

Ultraviolet astronomy focuses on the area of the spectrum that is beyond violet light—the shortest wavelengths the eye can see. The ultraviolet portion of the spectrum begins at a wavelength of 390 nanometers, ranges to the extreme ultraviolet at 90 nanometers, and merges into the X-ray portion of the spectrum at 10 nanometers.

While the visible portion of the spectrum can be observed from the surface of the Earth, observations at ultraviolet wavelengths must be done outside the Earth's atmosphere. Ultraviolet radiation is readily absorbed by gases, both in space and in the Earth's atmosphere. Only the longest wavelengths of ultraviolet light penetrate the atmosphere. It is this radiation that is responsible for the destructive tanning effects of the Sun on the skin. Screening effects of the atmosphere protect life on Earth from the more harmful, shorter-wavelength ultraviolet radiation. At the same time, the atmosphere prevents astronomers from easily collecting information about the universe from this important range of the spectrum.

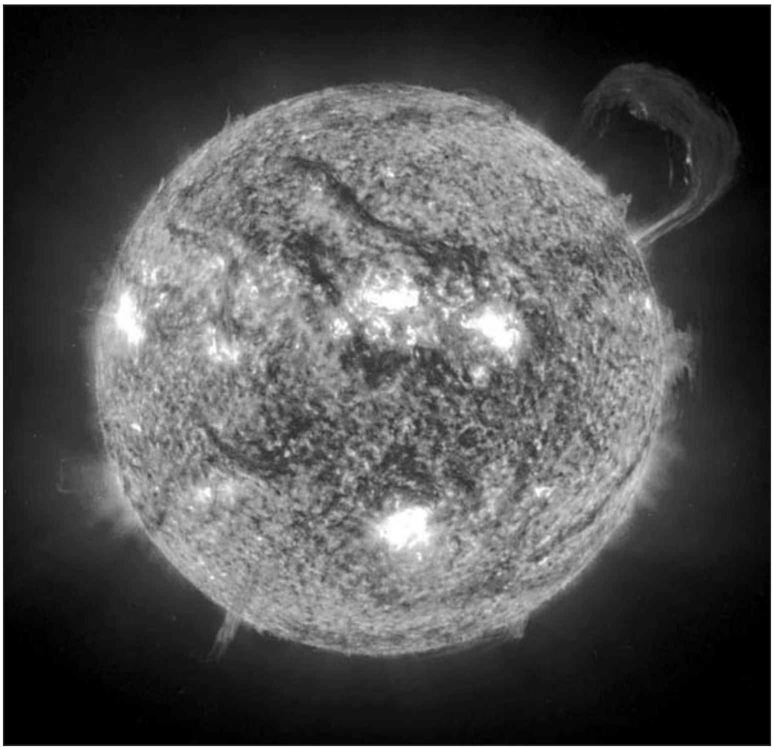
Practically every object in the universe emits some radiation at ultraviolet wavelengths. Any material that has a temperature between 10,000 and 1 million kelvins thermally emits most of its energy in the ultraviolet. This range of the spectrum is important to the knowledge of celestial objects, since the atmospheres of most stars, the surfaces of very massive stars, white dwarf stars, and regions of hot interstellar gas all fall within this temperature range. Every element emits and absorbs energy according to a characteristic pattern. By analyzing the pattern, or spectrogram, astronomers can determine the chemical composition of very distant objects. A spectrum is governed by the temperature, density, and chemical composition of the object emitting the energy, as well as how the energy has been altered by intervening processes en route to the instruments. In addition, the elements that are most abundant in the universe, such as hydrogen, helium, carbon, nitrogen, oxygen, and silicon, all have spectral features that are prominent in the ultraviolet. For this reason, ultraviolet astronomy can provide astronomers with important information about the universe.

Although instruments used in ultraviolet astronomy are designed to operate remotely, they can be very similar to optical instruments. Ordinary telescope mirrors will focus ultraviolet light. Electronic detectors record the image, or, in some cases, the image can be recorded on regular photographic film. A spectrograph can record such information by passing the radiation through a narrow slit and then through a prism, which separates the radiation into its component wavelengths. The result, a spectrogram, is then recorded on film.

The first ultraviolet telescopes were flown on high-altitude weather balloons. Later, they were launched on rockets, which raised them above the atmosphere for a few minutes at a time. The best way to gain access to ultraviolet information is to place an ultraviolet observatory in orbit. The first satellites to carry ultraviolet instruments were the Orbiting Astronomical Observatories (OAOs), a series of four identical satellites that carried different instrument packages to measure ultraviolet radiation from stars and interstellar gas. Only two of the four spacecraft proved functional. The first satellite failed after two days in orbit, and another failed to reach orbit. The second OAO achieved a general ultraviolet survey of the sky, discovering ultraviolet sources within the galaxy and measuring ultraviolet light from bright nearby galaxies. The final OAO, Copernicus, was the first to target specific ultraviolet sources. Carrying a 0.8-meter telescope, it was launched on August 21, 1972, and was functional for nine years. Copernicus took the first detailed look at objects in a wide range of the ultraviolet spectrum.

Detailed ultraviolet pictures of the Sun were produced with the OAO mission and later using the Solar Maximum Mission (SMM) satellite. Spectacular ultraviolet solar studies resulted from the American crewed space station Skylab, which was launched in 1973. Three crews of three astronauts inhabited the space station for a total of five and one-half months. Throughout this time, they kept a continuous surveillance of the Sun with the Apollo Telescope Mount (ATM). The ATM carried eight telescopes, which observed the Sun in wavelengths ranging from visible light through the ultraviolet and into the X-ray range.

In 1978, two years before the OAO Copernicus ceased operation, the International Ultraviolet Explorer (IUE) was launched. The IUE was a joint venture by the National Aeronautics and Space Administration (NASA), the European Space Agency (ESA), and the British Science Research Council (SRC). Although the IUE was equipped with a telescope smaller than that of Copernicus (only 45 centimeters), it carried more modern ultraviolet detectors and was able to observe much fainter stars over a broader range of wavelengths. The IUE was run much like a traditional observatory and, as such, was designed to be used by visiting scientists rather than by a select group of researchers. The IUE remained stationed in an orbit that is geosynchronous (remaining constantly over one area of Earth), about 36,000 kilometers above the Atlantic Ocean, where it remained in continuous contact with at least one of two ground stations



This 1999 image of the Sun was taken in 304-angstrom wavelengths by the Extreme Ultraviolet Imaging Telescope. (ESA/NASA/SOHO) in the United States and Europe. This was an improvement over earlier satellites, which could not remain in continuous contact with ground stations. Their observations were also limited by low orbits, with a large percentage of the field of view blocked by the Earth. The IUE was decommissioned on September 30, 1996. Sadly it was merely shut off due to budget concerns, even though this workhorse observatory was functioning nearly at full capability.

Astro 1 was a Spacelab mission launched aboard the space shuttle in early 1990; it carried three ultraviolet telescopes. For the duration of the mission, project scientists conducted observations in the ultraviolet and X-ray regions. Other important ultraviolet satellites have been the European TD-1 and the Soviet Astron satellite. The TD-1, launched in 1972, measured the magnitudes of more than thirty thousand stars in four different spectral regions and gathered ultraviolet spectra from more than one thousand stars. The Soviet Astron satellite, launched in 1983, was similar in size and scope to the OAO Copernicus.

Ultraviolet astronomy advanced greatly with observatories such as the Hubble Space Telescope (HST), launched from the space shuttle Discovery on mission STS-31 on April 25, 1990; the Extreme Ultraviolet Explorer (EUVE), launched on June 7, 1992; and the Far Ultraviolet Spectroscopic Explorer (FUSE), launched on June 24, 1999. The HST has proved to be a powerful ultraviolet instrument, making exciting discoveries in the far ultraviolet, despite preliminary optical difficulties, which were repaired in 1993 during space shuttle mission STS-61. Thereafter the HST returned amazing images of distant galaxies, interstellar dust, and other objects from deeper in space than had ever before been observed. The EUVE was designed to survey the cosmos for objects emitting very energetic short-wavelength radiation in order to discover many new objects, including perhaps ten times as many white dwarf stars as previously known. FUSE was designed to carry a 2-meter ultraviolet telescope that would investigate objects in the far ultraviolet and extreme ultraviolet. NASA's Wind mission was conceived to use multiple instruments, including ultraviolet instruments, to study the solar wind and sample plasma waves, energetic particles, and electric and magnetic fields. The European Space Agency's Lyman Observatory was designed to examine the dynamics of the Milky Way's halo as well as comets.

FUSE was designed for only a three-year primary mission but was still in reasonable operational condition in 2002, despite having difficulties with its reaction wheels (used for pointing the observatory at selected celestial targets). Budget issues almost resulted in cessation of FUSE observations, but the use of the orbital telescope was extended until July 12, 2007, when controllers lost the final reaction wheel. Despite efforts to restore the observatory, on September 7, 2007, FUSE investigations ceased. This relatively inexpensive Explorer-class satellite (also known as Explorer 77) led to more than four hundred published scientific papers and advanced the careers of many young astronomers specializing in ultraviolet astrophysics.

APPLICATIONS

Each energy region in the electromagnetic spectrum allows astronomers to "see" objects in a unique way. The more information that can be discovered about the nature of a celestial object in each of these energy areas, the more completely the object can be understood.

All objects known to exist in the universe—from comets and planets to stars, galaxies, and quasars—can be effectively studied in the ultraviolet range. Ultraviolet telescopes see the hottest stars of all; as a result, they tend to pick out the youngest star groups in the sky. Ultraviolet astronomy can thus focus on the youthful clusters of stars that lie close to regions of star birth. With this particular window, ultraviolet astronomy has been very useful for mapping regions of star formation, both in the Milky Way and in distant galaxies. Other galactic studies have shown that the Milky Way, as well as other galaxies, surrounded by a hot halo of gas.

There are excellent images of the Sun in the ultraviolet. Views of the ultraviolet Sun reveal different layers of its chromosphere, transition region, and lower corona. Bright, scintillating points of ultraviolet light in the Sun's atmosphere provide a measure of magnetic activity within the Sun, with perhaps even more accuracy than the sunspots that are seen on the visible photosphere. By observing other stars in the ultraviolet, astronomers have gained valuable knowledge about the nature of stars, which correlates with what is known about the Sun. It has been shown that many stars have hot outer atmospheres similar to the Sun. A new class of stars was discovered that had distinguishing characteristics visible only in ultraviolet light.

Ultraviolet astronomy has been very useful in studying binary star systems. A binary system is a pair of stars orbiting around their common center of mass. In a binary system, one star can be much brighter in optical wavelengths, and only a single spectrum can be observed. If the companion star is much hotter, however, it will dominate the spectrum of the system at ultraviolet wavelengths. A binary system gives astronomers a tool with which they can study the nature of these dimmer hot stars. Previously unobserved hot companions have been discovered in stars not suspected earlier of being binaries.

The supernova is another area in which ultraviolet astronomy can contribute significantly. A supernova is a stellar explosion in which all or most of the star's mass is expelled. Astronomers have studied the remains of supernova explosions as well as observing them in the beginning stages of development. Ultraviolet observations can determine the chemical composition of the layers of the star expelled by the explosion. Astronomers discovered that a nova explosion in Cygnus in 1978 produced much nitrogen, while the supernova that created the well-known Crab nebula threw out relatively small amounts of carbon. These facts are important clues to learning how new elements are formed, as well as in understanding the mechanisms behind supernovae.

The IUE observed Comet Kohoutek in 1976, finding a very bright image in ultraviolet wavelengths. Comet IRAS-Iraki-Alcock was observed in 1983 and Comet Halley was observed in 1986. In combination with observations at other wavelengths, it was found that the comets are similar in composition, suggesting that they have a similar origin.

Among IUE's legacy of ultraviolet investigations were 104,000 individual observations, the discovery of aurorae at Jupiter, the first determination of the water-loss rate in a comet, the production of the first orbital radial velocity curve for Wolf-Rayet stars, determination of the progenitor of Supernova 1987A, the first imaging of galactic halos, and the production of 44,000 stellar spectra per year, to name but a few. In a very real sense, IUE, more than anything before it, expanded ultraviolet astronomy from an interesting concept to a highly active and fruitful area of astrophysical observations.

Ultraviolet astronomy has confirmed some long-standing theories that previously lacked sufficient evidence. The theory of a "gravitational lens" had been predicted by the relativity theories of Albert Einstein but had never been supported by solid evidence. The gravitational lens is a process in which the gravitational field of a very massive object acts as a lens, bending the radiation from a more distant object behind it, distorting its image and often creating a double or multiple image. Observations by the IUE helped to indicate that such gravitational lenses exist. The first such lens system was discovered in 1979. Hubble routinely discovered new gravitational lensing objects.

A full list of science achievements of EUVE would be too long to include in this article. Among some of those discoveries were the production of an all-sky catalog containing a total of 801 ultraviolet-emitting objects (in wavelengths from 7 to 76 nanometers), participation in coordinated observations of numerous objects at different wavelengths across the electromagnetic spectrum, some of the first ultraviolet detections of extragalactic objects, measurements of quasi-periodic oscillations in dwarf novae, analyses of extreme ultraviolet spectral white dwarf star companions to main sequence stars, the detection of helium in a hot white dwarf, and the first extreme ultraviolet observations of the Coma Cluster.

Ultraviolet observations with the Hubble Space Telescope involved uses of the observatory's Goddard High Resolution Spectrograph (GHRS), Faint Object Camera (FOC), and Faint Object Spectrograph (FOS). GHRS and FOS were replaced during the second Hubble servicing mission (STS-82 in February, 1997). In their place were inserted the Space Telescope Imaging Spectrograph (STIS) and Near Infrared Camera and Multi-Object Spectrometer (NICMOS). During shuttle servicing mission 3B (STS-109 in March, 2002) the FOC was replaced by the Advanced Camera for Surveys (ACS). STIS suffered a malfunction in August, 2004, and portions of the electronics for ACS failed in 2006 and 2007. STIS and ACS were scheduled to be repaired during one final shuttle servicing mission, during which repairs would restore Hubble's capability to conduct ultraviolet astronomy. That shuttle mission was also intended to insert a new Wide Field Camera 3 and the Cosmic Origins Spectrograph to expand the observatory's capability to make ultraviolet measurements of objects at great distances (and hence tremendously early times in the cosmic past).

The July, 2008, edition of *Physics Today* presented research using ultraviolet data from both FUSE and HST that indicated that as much as 40 percent of the anticipated baryonic matter is missing from the portion of the universe along the line of sight to energetic objects such as quasars. The study involved collecting ultraviolet spectral signatures of quasars. What was anticipated was a typical quasar spectrum, which also incorporated into it absorption lines for the absorbing gas between the quasar source and the telescope in Earth orbit detecting that radiation. By analyzing the depth of absorption lines cutting into the quasar's spectral emissions, astronomers were able to calculate the density of the baryonic matter (originally formed within just a few minutes of the big bang) along the line of sight. It turned out to be too low. Obviously, more investigations of the intergalactic medium were necessary, or an adjustment of cosmological models would be in order.

FUSE provided insight into the abundance of deuterium in stars and studied a wide range of astrophysical objects, such as the intergalactic medium, cool stars, and galactic structures. By recording absorption and emission lines in the far-ultraviolet portion of the electromagnetic spectrum, FUSE increased our understanding of galactic, intergalactic, and extragalactic chemical processes.

Context

For thousands of years, the human eye was the only astronomical instrument. The eye evolved to be most sensitive to the range of visible light, the most abundant source of radiation at the surface of the Earth. Any celestial objects that were dim or emitted radiation at mostly nonoptical wavelengths remained invisible to the eye, which limited the range of information about the universe scientists could study.

The invention of the telescope radically altered astronomy, not only because of the fainter objects it allowed astronomers to see but also because it opened up the possibility that there was more to the universe than what the human eye was able to image. Astronomy improved dramatically over the next few centuries but remained optical. The first sign that there was another way to look at the universe with anything other than optical wavelengths came in 1800, when infrared radiation was discovered by Sir William Herschel, who placed a thermometer just outside the red range of the visible light separated by a prism.

The opening up of the wavelengths of electromagnetic energy got under way with the rapid growth of radio astronomy in the 1950's and 1960's and with the birth of the space program during the same period. The space program allowed ultraviolet astronomy to become an important new area of study. The potential value of ultraviolet observation from space was proposed to the U.S. Air Force in 1946 by the American astrophysicist Lyman Spitzer, Jr. With the establishment of NASA in 1958, the concept of placing orbiting observatories in Earth's orbit became a reality, and a series of orbiting observatories were launched over the next twenty years.

The result of more than forty years of observing in all ranges of the spectrum is that astronomers now have a more complete understanding of the processes occurring in the universe. Today's astronomers have taken images of the stars and galaxies that were unimaginable to the ancients, or even to the astronomers of a few decades past. Ultraviolet astronomy is now at a point where future missions will lose the "frontier" feel of the early missions, with increasingly complex and specialized missions. Even so, exciting new discoveries will continue to be made. In addition to new observations, research using information from the years of observations made by the satellites since the 1960's will allow astronomers to gain new insights into virtually every area of the universe.

Divonna Ogier

FURTHER READING

Arny, Thomas T. *Explorations: An Introduction to Astronomy*. 3d ed. New York: McGraw-Hill, 2003. A general astronomy text for the nonscientist. Includes an interactive CD-ROM and is updated with a Web site.

Barstow, Martin A., and Jay B. Holberg. *Extreme Ultraviolet Astrophysics*. Cambridge, England: Cambridge University Press, 2007. The universe in extreme ultraviolet (EUV) was revealed only when rockets in the late 1960's sent relatively primitive instruments capable of detecting EUV radiation briefly into space. Those initial investigations demonstrated the universe was rich in EUV-emitting sources. This work catalogs EUV sources, explains the cosmological importance of those sources, and describes the instrumentation that detected them.

Henbest, Nigel. *Mysteries of the Universe*. New York: Van Nostrand Reinhold, 1981. Explores the limits of what is known about the universe. Discusses theories about the origin of the solar system and the universe, exotic astronomy, and astronomy at invisible wavelengths.

Karttunen, H. P., et al., eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007. A well-used university textbook in introductory astronomy. Contains some calculus-based treatments for those who need a more advanced textbook than the standard introductory work. Suitable for an audience with varied science and mathematical backgrounds. Covers all topics from solar-system objects to cosmology.

Marten, Michael, and John Chesterman. *The Radiant Universe*. New York: Macmillan, 1980. An overview of imaging that is not accessible in visible wavelengths. Discusses electronic processing as well as infrared and ultraviolet wavelength imaging. Beautiful pictures, along with easy-to-read and informative text, somewhat compensate for the age of this text.

Time-Life Books. *The New Astronomy*. Alexandria, Va.: Author, 1989. One volume of a series examining different aspects of the universe. Comprehensively covers all invisible astronomies, including high-energy astronomy and imaging techniques. Heavily illustrated and suitable for the general reader with an interest in astronomy.

See also: Archaeoastronomy; Coordinate Systems; Earth System Science; Gravity Measurement; Hertzsprung-Russell Diagram; Infrared Astronomy; Neutrino Astronomy; Optical Astronomy; Radio Astronomy; Solar Corona; Telescopes: Ground-Based; Telescopes: Space-Based; X-Ray and Gamma-Ray Astronomy.

Universe: Evolution

Category: The Cosmological Context

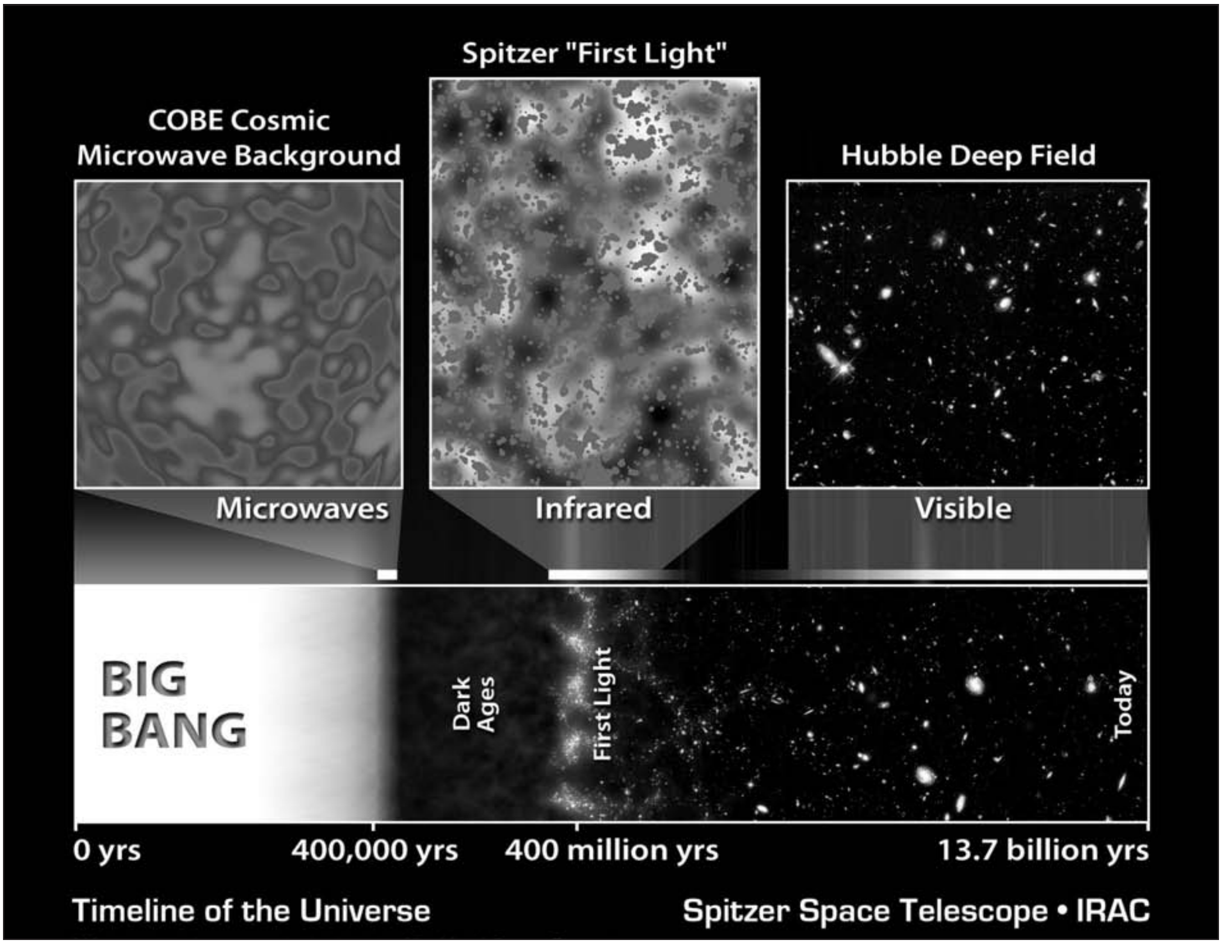
About 13 to 14 billion years ago, the big bang occurred, creating the space, time, matter, and energy of our universe. Since then, space has expanded, galaxies of stars have formed in that space, and the stars in the galaxies have evolved. The evidence indicates that the universe will continue to expand forever. Indeed, the expansion appears to be accelerating, due to some unknown cause referred to as "dark energy."

OVERVIEW

Evidence we have from observing radiation in all ranges of the electromagnetic spectrum—particularly the infrared—has led astrophysicists and cosmologists to conclude that, approximately 13 to 14 billion years ago, an explosive event dubbed the big bang created space and time, matter and energy from an unimaginably hot, dense state whose origin is unknown. Ever since the big bang, the universe has been expanding and evolving. Events following the big bang have been reconstructed in great detail by using known physical laws to predict the behavior of matter and energy as the universe expands.

As the universe expanded, its temperature and the density of both matter and electromagnetic radiation all decreased. In the early, very hot, very dense universe, cosmologists think that all four fundamental forces of nature—gravity, the strong nuclear force, the weak nuclear force, and the electromagnetic force—were unified as one force, indistinguishable because as yet unseparated from each other. Models that would successfully combine these four forces are called theories of everything (or TOEs), but currently there is no workable TOE.

All TOEs, however, assume that the four forces were unified only at the extremely high energies, corresponding to extremely high temperatures, that characterized the first few seconds of the universe. Sometime between 10 and 43 seconds after the big bang (called the Planck time), the temperature dropped to about 10^{32} kelvins, and gravity separated ("froze out") from the other forces. After about 10^{-35} second, the temperature dropped to about 10^{28} kelvins, and the strong nuclear force separated from the remaining two forces. The "freeze-out" of the strong nuclear force may have been what initiated a period of rapid inflation, in which the universe expanded exponentially, increasing in size by a factor of about 10^{50} in the next 10^{-32} second. This rapid inflation explains why distant regions of the present universe appear so simi-



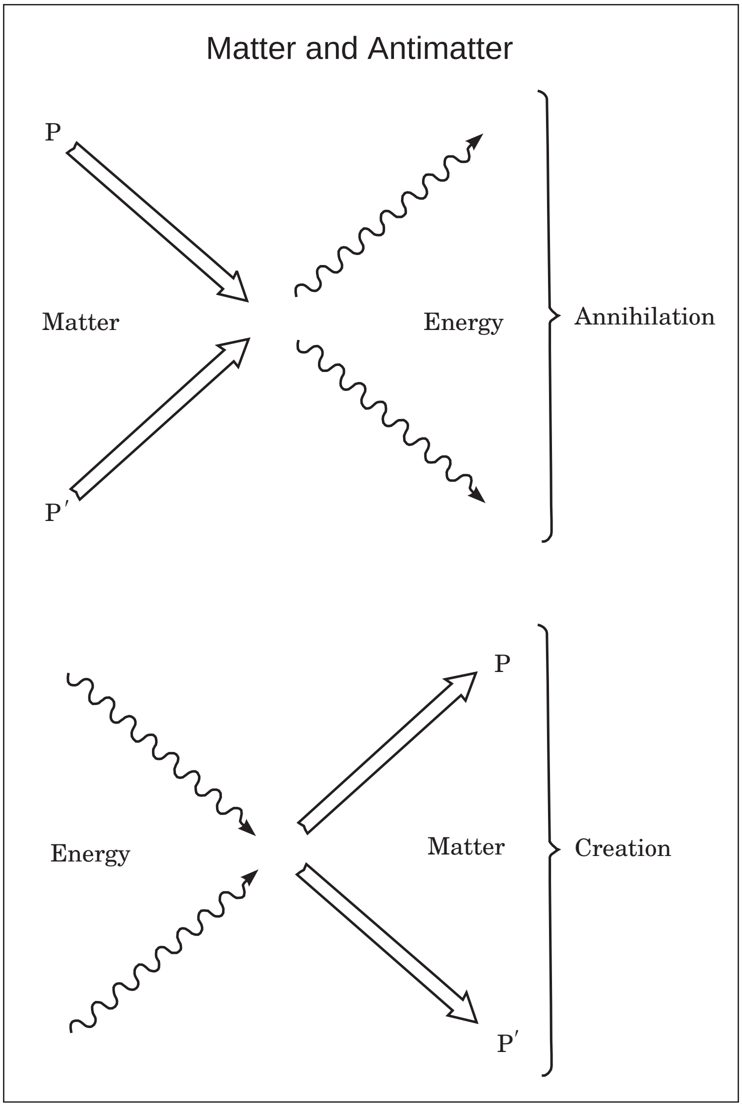
A time line of the history of the universe, from the big bang on, showing some of the evidence that has been collected by space telescopes capturing images in the microwave, infrared, and optical ranges. (NASA/JPL-Caltech/A. Kashlinsky, GSFC) lar, and why space on the large scale is so nearly flat. Then the universe resumed the slower expansion it had been undergoing before inflation occurred. After about 10^{-10} second, the temperature had dropped to 10^{15} kelvins, and finally the weak nuclear and electromagnetic forces assumed their separate characteristics.

The early universe was very hot and was filled with high-energy gamma-ray photons. When two gamma rays with sufficient energy collided, their energy could be converted to mass, as described by Einstein's famous equation, $E = mc^2$, which says that energy, E, and mass, m, are equivalent and related by the speed of light, c, squared. The mass appeared as a particle-antiparticle pair in a process called pair production. When a particle and its antiparticle encountered each other, they would mutually destroy each other in a process called annihilation, in which their mass would be converted back into energy as a pair of gamma-ray photons. The rates of pair production and annihilation were equal, and matter and radiation were in a state of thermal equilibrium. However, as the universe continued to expand, the temperature dropped to the point that the photons no longer had enough energy to produce specific particle-antiparticle pairs. Once the temperature had dropped below the threshold temperature for a particular type of pair production, no more of those particle-antiparticle pairs were formed, and those that had formed previously quickly annihilated each other.

The threshold temperature for proton-antiproton pair production is about 10^{13} kelvins; this temperature was reached when the universe was about 10^{-4} second old, after which no more proton-antiproton pairs were produced, and those protons and antiprotons that had been produced previously annihilated each other. The threshold temperature for electron-positron (another name for an antielectron) pair production is about 6×10^9 kelvins; that temperature was reached after a few seconds, after which no more electron-positron pairs were produced, and those electrons and positrons that had been produced previously annihilated each other. If exactly equal numbers of particles and antiparticles had been created, they all would have annihilated each other, and there would be no matter or antimatter in the universe today. However, we live today in a universe composed predominantly of matter. Consequently, a slight excess of particles over antiparticles must have been created, by about one part in a billion; they survived and constitute the matter in the present universe.

Electrons could combine with protons to form neutrons, and protons and neutrons could combine to form nuclei of deuterium (also called heavy hydrogen), each deuterium nucleus consisting of one proton and one neutron held together by the strong nuclear force. High-energy gamma rays could break deuterium nuclei back into protons and neutrons as fast as they had formed, but by about three minutes after the big bang, the temperature had dropped to about one billion kelvins, and photons no longer had enough energy to break up deuterium nuclei. This began the time of nucleogenesis, when deuterium nuclei could fuse into helium nuclei and even form some lithium and beryllium nuclei. However, after about 15 minutes, the temperature had dropped to a few hundred million kelvins, and the nuclei no longer were moving fast enough to overcome their electrical repulsion. The nucleosynthesis of heavier elements by fusion would have to wait till much later, when it would occur in stars. This established the overall composition of the universe today—about one atom of helium for every ten atoms of hydrogen, with only very small amounts of all the other chemical elements.

In the early universe, the density of electromagnetic radiation was greater than the density of matter, but as the universe expanded the radiation density decreased more quickly than the matter density. Several thousand years after the big bang, at a time called the crossover time, the density of radiation and matter were equal. From that time on, the universe has been dominated by matter. At the high temperatures of the early universe, the matter was ionized, meaning it consisted of free electrons and bare atomic nuclei. Free electrons are very effective at scattering photons of electromagnetic radiation, so the early universe was opaque; photons could not travel far before encountering free electrons and being scattered in new directions. About 300,000 to 400,000 years after the big bang, the temperature dropped to about 3,000 kelvins, and the electrons could combine with nuclei to form neutral atoms. Neutral atoms are able to absorb only certain specific photon energies, so the universe became transparent to most photons, and matter and radiation decoupled. The photons could travel freely through the universe, and today we observe them as the greatly redshifted cosmic microwave background radiation.



An elementary particle (P) and an antiparticle (P′) may collide, annihilating each other and releasing energy (wavy arrows). The reverse may also occur, creating matter. Such collisions are relatively rare at the present stage of the universe's evolution.

After matter and radiation decoupled, small fluctuations in the distribution of matter started to grow; slightly denser regions gravitationally attracted matter from surrounding areas and became denser still. Within the first billion years, they developed into small protogalaxies or pregalactic fragments in which the first stars formed by gravitational contraction of clouds of gas. Through mergers, the protogalaxies formed larger systems of stars, the galaxies that make up the universe today. Within the galaxies, stars continue to form by the gravitational contraction of gas clouds, and at least some stars develop families of planets as a by-product of their own formation. It is stars that synthesize the heavier chemical elements. Some of the elements are formed by nuclear fusion processes during the active, energy-producing lives of the stars. Others are formed when stars explode as supernovae. Whether stars end their energy-producing lives violently as supernovae or more quietly, they expel some or most of their mass, and this disperses the heavier elements into the interstellar material, enriching the clouds of gas and dust from which new stars form.

The Milky Way galaxy formed more than 12 billion years ago. The Sun and its solar system formed from a cloud of gas and dust about 4.5 billion years ago. Without the nucleosynthesis of heavier elements by earlier generations of stars, there would be no carbon, oxygen, silicon, iron, or any of the many other elements needed to form a rocky planet like Earth and the life it supports.

Observations of distant galaxies indicate that the expansion of the universe is accelerating. The cause is unknown, but it is has been given the name "dark energy." If the expansion continues to accelerate, the distances between galaxies will grow ever greater. Eventually, all the matter in the galaxies will be processed into stars, the stars will all run out of energy and die, and the universe will grow dark and cold.

METHODS OF STUDY

An expanding universe is predicted by Albert Einstein's general theory of relativity. This conclusion was arrived at independently by Alexander Friedmann in 1922 and Georges Lemaître in 1927 from solutions they found to the field equations of general relativity applied to the structure of the universe. In 1929, Edwin Powell Hubble showed that the universe actually is (or at least appears to be) expanding when he discovered that the distances of thirty-one galaxies were correlated with the redshifts of their spectra. The cosmological explanation for this relationship is that, from the perspective of observers on Earth, the expansion of space stretches the wavelengths of electromagnetic radiation as it travels through space from the source to the observer. The greater the distance between source and observer, the longer it takes electromagnetic radiation to travel the distance, the longer the universe has been expanding, and the more the wavelengths are stretched. The long-wavelength end of the visible light spectrum is the red end, so the expansion of space shifts visible light to longer, redder wavelengths. The term "redshift" refers to a shift toward longer wavelengths of a photon of light in any portion of the electromagnetic spectrum, whether visible light or not, hence a shift toward the lower energy of that photon.

The cosmic scale factor $R(t)$ is defined to be a measure of how the universe changes in size as a function of time. Changes in the cosmic scale factor can be determined directly from the spectral redshifts. The amount the wavelengths are lengthened tells cosmologists how much the universe has expanded since the light was emitted; for example, if the features of some spectrum all have double their expected wavelengths, then the size of the universe and the cosmic scale factor have both doubled since the light was emitted; to put it another way, the universe and cosmic scale factor then both were half as big as they are now.

Lemaître, in 1927, was the first to propose that the expansion of the universe began from a compact, dense initial state—the "primeval atom" as he called it—which "fissioned" into all atoms in the universe today. Although wrong in the details, Lemaître's basic idea of expansion and evolution from a compact initial state has been developed into the big bang model of modern cosmology. Events following the big bang have been reconstructed in great detail by using known physical laws to predict the behavior of matter and energy as the universe expands.

Space is filled with electromagnetic radiation that has been traveling freely since the universe became transparent a few hundred thousand years after the big bang. Its serendipitous discovery by Arno Penzias and Robert Wilson in 1965 was a significant confirmation of a hot, dense big bang origin for the universe. This is thermal blackbody radiation, so the wavelength at which it is "brightest" is inversely proportional to the temperature. As the universe has expanded, the wavelength at which it is brightest has increased linearly with the cosmic scale factor R. Thus the temperature T decreases as the inverse of the cosmic scale factor R. Since the temperature of the cosmic background radiation now is about 3 kelvins and the temperature at which the universe became transparent was about 3,000 kelvins, the universe has expanded by a factor of about 1,000 since then.

The expansion of the universe causes the density of both matter and radiation to decrease with time. Since volume increases as R^3, the density of matter decreases as the inverse of R^3. However, the density of the energy of electromagnetic radiation decreases as the inverse of R^4. This is because the number of photons of electromagnetic radiation per volume of space decreases as the inverse of R^3, and the energy of each photon decreases as the inverse of R; the wavelength associated with each photon increases linearly with R, and wavelength and photon energy are inversely related. Thus the density of electromagnetic energy decreases as the inverse of R^3 times the inverse of R, which equals the inverse of R^4. This means that the density of radiation energy decreases more rapidly than matter density as the universe expands. Currently the density of matter is several thousand times the density of electromagnetic energy, but at an earlier time, when the universe and the cosmic scale factor were several thousand times smaller than they are now, the density of matter and electromagnetic energy were equal. The time when the densities of matter and electromagnetic radiation were equal is called the crossover time; before that, the density of electromagnetic radiation was greater than the density of matter. Current estimates of the crossover time place it several thousand years after the big bang.

The discovery that the expansion of the universe is accelerating was completely unexpected; in fact, it was discovered during an attempt to measure how much the expansion was decelerating. The models of the universe derived from the simplest form of general relativity all predict the expansion should be slowing. If it were slowing only a little, then the universe would be open and the expansion would continue forever. If it were slowing enough, then eventually the expansion would stop and the universe would begin to contract back on itself. It was expected that measuring the redshifts of very distant galaxies would show how much faster the expansion was in the long-ago past, when the light we now receive left those galaxies. However, the observational evidence indicates the expansion was slower in the past, meaning the expansion has been accelerating.

CONTEXT

Throughout history, humankind has wondered about the origin and fate of the Earth, its life, and the universe. This desire to understand our origins has led nearly every culture to form some kind of "creation myth." In Western culture, many religious and philosophical beliefs about the origin of the universe can be traced back thousands of years to the creation myths of the Middle East. Although science cannot explain the origin of what Lemaître called the "primeval atom," evidence from observations of phenomena billions of light-years old has now provided more definitive answers to many questions that humans have pondered for thousands of years.

Perhaps the most crucial of these questions is how the universe formed. Physics, coupled with astronomical observations, has helped us work out the events and processes that likely occurred in the aftermath of the big bang. Our models of the evolution of the universe have profound implications for understanding life on Earth. The universe seems "fine-tuned" for the existence of life as we know it. If the physical laws and constants were changed slightly, the universe would be a very different place, and life as we know it could not exist. Some scientists explain this by invoking what is called the anthropic principle: The universe has to be the way it is because we are here; if conditions did not permit the development of life, we would not be here to speculate about it. Others argue that the odds are too great against it just being chance that the universe is the way it is, and suggest that it may have been deliberately designed that way. Still others speculate that our universe is just one of many universes, each with its own physical laws and constants; we live in the one in which life as we know it is possible.

Another aspect of the evolution of the universe with profound implications is its future and ultimate fate. If the expansion continues to accelerate, eventually, some billions of years from now, all matter will be processed into stars, all stars will run out of energy, and the universe will grow cold and dark.

Michael L. McKinney and Richard R. Erickson

FURTHER READING

Belusevic, Radoje. *Relativity, Astrophysics, and Cosmology*. Weinheim, Germany: Wiley-VCH, 2008. Addresses the interrelationship of the now entwined disciplines of astrophysics, particle physics, cosmology, and relativity theory, which have combined to form advanced theories of the origin and evolution of the universe.

Bennett, Jeffrey, et al. *The Cosmic Perspective*. 5th ed. San Francisco: Pearson/Addison-Wesley, 2008. Structured around a set of two-page figures dubbed "cosmic contexts," this interactive resource uses zoom-in illustrations to orient students to various images of the universe in relation to one another, while covering all the basics of cosmology. For general and introductory audiences. Chapters include “Light and Matter: Reading Messages from the Cosmos,” “A Universe of Galaxies,” “Dark Matter,” “Dark Energy,” and “The Fate of the Universe.”

Drexler, Jerome. *Discovering Postmodern Cosmology: Discoveries in Dark Matter, Cosmic Web, Big Bang, Inflation, Cosmic Rays, Dark Energy, Accelerating Cosmos*. Boca Raton, Fla.: Universal, 2008. Cosmologist Drexler proposes a plausible and unique correlation of the seven mysterious areas of cosmological research listed in the book’s subtitle. For all open-minded audiences.

Duncan, Todd, and Craig Tyler. *Your Cosmic Context: An Introduction to Modern Cosmology*. San Francisco: Pearson/Addison-Wesley, 2009. An introductory texbook for studies in modern cosmology that engages students by relating cosmological concepts to their own lives.

Ferreira, Pedro G. *The State of the Universe: A Primer in Modern Cosmology*. London: Phoenix, 2007. Oxford lecturer Ferreira presents a history of cosmology, examining the complexities that concepts such as dark matter and dark energy have imposed on a once “simple” Einsteinian universe ruled by relativity.

Gasperini, Maurizio. *The Universe Before the Big Bang: Cosmology and String Theory*. Berlin: Springer, 2008. A fascinating exploration of what might have happened prior to the explosion that formed the universe, which looks to string theory and other modern mathematical models to postulate that the universe was not born with the big bang but rather was well advanced in its overall evolution. Presented with nontechnical language for nonspecialists.

Lemoine, M., J. Martin, and P. Peter, eds. *Inflationary Cosmology*. New York: Springer, 2008. This collection of papers, by both venerable and younger astrophysicists and cosmologists, focuses on that period during which the early universe expanded at a greatly accelerated rate before settling down to a slower rate of expansion. Presents several different scenarios.

Liddle, Andrew, and Jon Loveday. *The Oxford Companion to Cosmology*. New York: Oxford University Press, 2008. An indispensable A-Z reference, consisting of more than 350 entries from antimatter to WIMPs, as well as individual physicists and other scientists who have advanced the field. Heavily illustrated with almost two hundred halftones and diagrams; includes cross-references and Web links.

North, John. *Cosmos: An Illustrated History of Astronomy and Cosmology*. Chicago: University of Chicago Press, 2008. Emphasizes the astrophysics of cosmology, with a focus on physical and mathematical concepts that are key to understanding classical field theory. Also describes experimental techniques and results. Technical.

Sagan, Carl. *Cosmos*. New York: Random House, 1980. One of the most popular science books of all time. It is superbly written and illustrated by a famous astronomer and is based on the popular television series of the same name. Fun and easy reading.

Schneider, Peter. *Extragalactic Astronomy and Cosmology: An Introduction*. New York: Springer, 2006. A textbook focusing on galaxies, clusters, and superclusters, beginning with the Milky Way. Covers their evolution, formation, and distribution, supported by beautiful color illustrations.

Weinberg, Steven. *Cosmology*. New York: Oxford University Press, 2008. The Nobel physics laureate presents the subject in two parts: one covering the isotropic and homogeneous “average” universe; the second, departures from the average universe. Provides detailed coverage of recombination, microwave background polarization, leptogenesis, gravitational lensing, structure formation, and multifield inflation. Includes mathematical calculations. Appendixes review general relativity, the Boltzmann equation, and sample problems. For advanced students and professionals.

Wudka, Jose. *Space-Time, Relativity, and Cosmology*. New York: Cambridge University Press, 2006. A history of relativistic cosmology from ancient times to Einstein. The nonmathematical approach emphasizes concepts over calculations yet explain the ideas clearly, applying them to research topics in

cosmology. Designed for students from high school through college.

See also: Big Bang; Cosmic Rays; Cosmology; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; General Relativity; Interstellar Clouds and the Interstellar Medium; Milky Way; Novae, Bursters, and X-Ray Sources; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Universe: Expansion; Universe: Structure.

Universe: Expansion

Category: The Cosmological Context

The universe is expanding, with the most distant galaxies receding at the greatest speeds. The rate of expansion and any change in the rate of expansion are critical data that provide constraints on the origin and ultimate future of the universe.

OVERVIEW

Galaxies are vast collections of millions to trillions of stars, all gravitationally bound together; our Sun and solar system are part of the Milky Way galaxy, which contains several hundred billion stars. There are billions of galaxies, and they are gravitationally bound together in galaxy clusters that contain between a few tens to a few thousand galaxies; our Milky Way galaxy is part of the cluster named the Local Group which contains over 40 known members, almost all of which are much smaller than our Milky Way galaxy.

Beginning in 1912, Vesto Melvin Slipher used the 24-inch refracting telescope at Lowell Observatory in Arizona to obtain spectra of objects called spiral nebulae. (At that time, it was not known that they actually were galaxies, although some astronomers thought they might be.) He found that, although a few of them had spectra that were shifted toward shorter, bluer wavelengths, most of their spectra were shifted toward longer, redder wavelengths, and some of the redshifts were surprisingly large. If the spectral shifts were Doppler in origin (produced by motion of the spiral nebulae along the line of sight), then most of the spiral nebulae were moving away from us at high speed.

In the 1920's, using the Mount Wilson Observatory's 100-inch reflecting telescope (then the largest in the world), Edwin Powell Hubble succeeded in resolving individual stars in some of the spiral nebulae, including M31—the spiral nebula in Andromeda—and NGC 6822. He found that some of the stars were Cepheid variables. Earlier, Henrietta Swan Leavitt at Harvard had discovered that the average luminosity of a Cepheid variable is related to its period of light variation. The bigger the Cepheid variable, the longer it takes to pulsate and thus the longer its period of light variation. Also, the bigger the star, the brighter its average luminosity. Using the periods of the Cepheid variables, Hubble could determine their luminosities, or real brightnesses. By comparing their real brightnesses with their apparent brightnesses, Hubble was able to determine their distances, which placed them, and the spiral nebulae that contained them, far outside the Milky Way galaxy. This showed that spiral nebulae really were spiral galaxies, similar to the Milky Way galaxy, as had been suggested by the philosopher Immanuel Kant and others as early as 1755.

In 1929, Hubble, assisted by Milton Humason, announced a correlation between the distances and spectral shifts of about thirty galaxies. Only a few of the nearer galaxies had blueshifted spectra; most had redshifted spectra with the size of the redshift proportional to distance (a relation now known as Hubble's law). If the redshift is due to motion away from us, then the farther away a galaxy is from us, the faster it is moving away. Hubble and Humason discovered that the universe appears to be expanding, with the distances between galaxies continuing to grow larger.

Actually, an expanding universe had been predicted earlier. In 1922, the Russian mathematical physicist Alexander Alexandrovich Friedmann and, indepedently in 1927, the Belgian priest and cosmologist Georges Lemaître had found two classes of solutions to the field equations of Albert Einstein's general theory of relativity. In one type of solution, called open, the universe continues to expand forever. In the other type of solution, called closed, the universe will expand to some maximum size and then contract. In 1932, Einstein and the Dutch astronomer Willem de Sitter found a third solution, called flat or critical, in which the universe expands forever, but just barely; it is the boundary case between the other two.

Lemaître was the first to propose that the expanding universe had its origin in a small, super-dense state—the "primeval atom." Later, this idea was expanded into a sort of primordial "explosion" that created space and time, matter and energy. In the 1950's, it came to be called the big bang. In all such big bang models, the universe evolves; as the distance between galaxies increases, the density of matter and the density of radiation decrease.

An alternative, the steady state theory, was presented by the British cosmologists Hermann Bondi, Tom Gold, and Fred Hoyle in 1948. In this model, as the universe expands, new matter is spontaneously created in the space between the galaxies at just the right rate to keep the average density of matter constant. Such a universe has neither beginning nor end. Both the steady state and big bang models had supporters throughout the 1950's, but by the mid- to late 1960's, observational evidence mounted that the universe was expanding from an early high-density state, and the steady state theory gradually fell out of favor.

A simple estimate of the time since the big bang can be found from Hubble's law. Since the speed of recession is proportional to distance, then at some time in the past all space was infinitesimally small. The relation of the slope of the speed v versus distance r is called the Hubble constant H, and $v = Hr$. The reciprocal of the Hubble constant, 1/H, is called the Hubble time or Hubble age, tH: $t\mathrm{H} = 1/\mathrm{H}$. If nothing has sped up or slowed down the expansion, the Hubble age is the time since the expansion began. Most recent determinations of H put it in the range of about 70 to 75 kilometers per second per million parsecs, which gives a Hubble age of 13 to 14 billion years.

However, it was expected that the expansion should slow down or decelerate with time, due to the gravity of all the mass in the universe. The only question was whether the expansion would slow down just a little and the universe would keep expanding forever, or whether the expansion would slow down enough to stop, after which the universe would begin to contract. In any case, the actual age of the universe would

Fred Hoyle and the Steady State Theory

Troubled by the problems presented by George Gamow's early theory of the big bang, Fred Hoyle and the University of Cambridge developed an alternative and well-respected proposal: the "steady state" theory.

Ironically, the term "big bang" had been coined by its main adversary—Hoyle himself—during one of his series of BBC radio talks. Hoyle used the term to belittle Gamow's theory. Hoyle favored a different view: that the universe, although currently expanding, was infinitely old and in the long term existed in a steady state. Galaxies were not receding from each other as the aftermath of a primordial explosion (which defenders of the big bang held). Rather, space was being created between galaxies at a constant rate, and hydrogen was being created to fill that space, coalescing into nebular clouds that then formed young stars and galaxies among the old.

The problem with this theory was that it contradicted the law of the conservation of matter: namely, that matter could neither be created nor be destroyed without being converted into energy. In the 1950's, the discovery of radio galaxies by Sir Martin Ryle revealed that galaxies had evolved billions of years ago, supporting the big bang theory.

Once Arno Penzias and Robert Wilson discovered the cosmic microwave background radiation, Hoyle's steady state theory was largely abandoned in favor of the theory he himself had named: the big bang. Although Hoyle revised his theory to account for the background radiation, his once dominant view of the universe was out of favor. Hoyle, however, remained philosophical to the end: "The Universe eventually has its way over the prejudices of men, and I optimistically think it will do so again."

be less than the Hubble age; if the universe were slowing down so much that eventually it would stop and then begin to contract, the actual age would be less than two-thirds the Hubble age. The difference in the cases depends, in part, on the density of matter in the universe. If the density were great enough, gravity would be strong enough to halt the expansion and cause the universe to contract; otherwise, the universe would continue to expand forever.

There are several observational tests for the ultimate fate of the universe. Attempts at measuring the density of matter that could be observed found much too low a density (by about a factor of 30) to stop the expansion, implying the universe would easily expand forever. On the other hand, the size of small fluctuations in the cosmic background radiation indicated the universe has a geometry that is very nearly flat, which is the geometry of the borderline case in which the universe just barely expands forever. Beginning in the 1990's, two independent research groups tried to detect a change in the rate of expansion, the presumed deceleration, by measuring the redshifts of very distant galaxies, and using Type Ia supernovae in them as "standard candles" to determine their distances. Contrary to all expectations, both groups found that the expansion has not been slowing down, but instead has been speeding up. This acceleration of the expansion has been attributed to some unknown mechanism called "dark energy," which seems to account for about 70 to 75 percent of all the matter and energy in the universe. If the acceleration continues, then the distance between clusters of galaxies will continue to grow at an increasing rate, and the universe will appear more and more empty.

METHODS OF STUDY

The expansion of the universe was initially predicted in the early 1900's by solutions to the field equations of Einstein's general theory of relativity. The observational demonstration of it comes from the redshift of galaxy spectra, which is interpreted to mean that galaxies are receding from each other. However, the expansion of the universe is not occurring because the galaxies are moving through space; the galaxies are moving with space as space itself expands. Thus the redshift is not due to the Doppler effect, which is produced by the motion of objects through space; instead, the redshift is termed cosmological, meaning it is due to the expansion of space itself. This might seem to be just a semantic difference, but there is a fundamental physical and mathematical difference. As space expands, carrying the galaxies with it, the wavelengths of electromagnetic radiation are stretched by the expansion of space. (Think of the wavelengths being drawn on the surface of a balloon that is being inflated. As the balloon expands, the wavelengths get longer.) The equation converting redshift into recessional speed is not the same for Doppler and cosmological redshifts. At small redshifts and speeds slow compared to the speed of light, both equations reduce to the same approximate form. However, for large redshifts and large speeds, the equations are significantly different. Conceptually, the two types of cause for the redshift are completely different, even though both involve objects receding from each other.

Since all forms of electromagnetic radiation (including visible light) have the same speed in vacuum, about 300,000 kilometers per second, astronomers can look back in time simply by observing objects at greater distances. The farther away something is, the longer it has taken electromagnetic radiation to reach us. Therefore, light from distant galaxies is light from the distant past. This is why cosmology generally requires the use of the largest telescopes, including both ground-based and space telescopes, so as to be able to observe objects that appear very faint because they are very far away. Observing galaxies at greater distances provides views of what the universe was like in the past and what it was doing then. This is how astronomers have discovered that the expansion was not occurring as rapidly in the past as it is now, meaning that the expansion is accelerating.

The greatest uncertainty about observing objects at large distances is determining just how large the distances are. Very large distances generally are measured by techniques called standard candle methods. A "standard candle" is an object whose luminosity or real brightness is known more or less accurately by some means. The known (or estimated) real brightness is compared to the measured apparent brightness to yield the distance.

Hubble and countless others since then have employed Cepheid variables as standard candles to determine galaxy distances. Cepheid variables are supergiant stars, thousands of times more luminous than our Sun. The period-luminosity relation for Cepheids discovered by Henrietta Leavitt more than a century ago is used to determine the average real brightness by measuring the period of light variation. However, even though Cepheids are intrinsically bright, they can be detected out to distances on the order of only 100 million light-years.

To detect changes in the expansion rate of the universe requires comparing the expansion long ago with the expansion now. Measuring the expansion long ago means observing galaxies at very great distances, out to billions of light-years. Type Ia supernovae are the standard candles used for such large distances. A Type Ia supernova is a white dwarf star that explodes violently when it acquires enough matter from a nearby companion star to push its mass over the Chandrasekhar limit of 1.4 solar masses. When this happens, it becomes about 3 billion times brighter than the Sun and can be seen at distances of billions of light-years.

Galaxies at such large distances, as determined by Type Ia supernovae in them, have had the redshifts of their spectra measured. It was expected that these data would show the expansion of the universe in the past, when the light we now receive left those galaxies was greater than it is now; just how much greater would show how much the expansion has slowed down. Contrary to expectations, however, the data show that the expansion in the past was slower than it is now, indicating that the expansion of the universe is accelerating.

CONTEXT

The discovery of the expansion of the universe forced a revolutionary change in ideas about the universe. Before then, the origin of the universe was a subject for metaphysics or theology. The expansion discovered by Hubble and Humason and the moment of creation first proposed by Lemaître and later called the big bang moved such questions into the realm of science.

The redshifted spectra of galaxies showed that the universe was expanding. The relation between redshift and distance, Hubble's law, allowed astronomers to estimate the age of the universe. The finite speed of light means that looking out to greater distances is equivalent to looking farther back in time. A major factor spurring the building of larger telescopes has been the desire to observe fainter objects at larger distances and hence at earlier times closer to the origin of the universe. Furthermore, the development of new technology and observing techniques has permitted astronomers to study the universe and its origin not only with visible light but also over the radio, infrared, ultraviolet, X-ray, and gamma-ray parts of the electromagnetic spectrum. All this has revealed more information not only about the origin of the universe but also about its future.

During most of the twentieth century, a major question about the expansion of the universe was whether the universe would continue to expand forever or would someday stop and eventually collapse. Astronomers needed observations of more distant galaxies to determine how much the expansion of the universe is slowing. However, such measurements made in the 1990's showed the expansion is not decelerating but accelerating. This has led to new theories and models about what has come to be called dark energy, the name given to whatever might be making the expansion accelerate.

Questions about the origin and expansion of the universe bring together general relativity and quantum mechanics, but as they stand now, these two theories are not fully compatible. Also, it is hypothesized that at the very high temperatures and energies in the first moments after the big bang, all four fundamental forces—gravity, strong nuclear, weak nuclear, and electromagnetic—were unified as one; as the universe expanded and cooled, the forces separated. Attempts to unify the strong nuclear, weak nuclear, and electromagnetic forces are called grand unified theories (GUTs). Theories that try to unify gravity with the other three forces are called theories of everything (TOEs). The details of such unified theories have yet to be worked out. Furthermore, new observations made with new observing techniques may result in quite unexpected discoveries, leading again to new ideas about the universe.

Pamela R. Justice and Richard R. Erickson

FURTHER READING

Belusevic, Radoje. *Relativity, Astrophysics, and Cosmology*. Weinheim, Germany: Wiley-VCH, 2008. Addresses the interrelationship of the now entwined disciplines of astrophysics, particle physics, cosmology, and relativity theory, which have combined to form advanced theories of the origin and evolution of the universe.

Bennett, Jeffrey, et al. *The Cosmic Perspective*. 5th ed. San Francisco: Pearson/Addison-Wesley, 2008. Structured around a set of two-page figures dubbed "cosmic contexts," this interactive resource uses zoom-in illustrations to orient students to various images of the universe in relation to one another, while covering all the basics of cosmology. For general and introductory audiences. Chapters include "Light and Matter: Reading Messages from the Cosmos," "A Universe of Galaxies," "Dark Matter," "Dark Energy," and "The Fate of the Universe."

Drexler, Jerome. *Discovering Postmodern Cosmology: Discoveries in Dark Matter, Cosmic Web, Big Bang, Inflation, Cosmic Rays, Dark Energy, Accelerating Cosmos*. Boca Raton, Fla.: Universal, 2008. Cosmologist Drexler proposes a plausible and unique correlation of the seven mysterious areas of cosmological research listed in the book's subtitle. For all open-minded audiences.

Duncan, Todd, and Craig Tyler. *Your Cosmic Context: An Introduction to Modern Cosmology*. San Francisco: Pearson/Addison-Wesley, 2009. An introductory texbook for studies in modern cosmology that engages students by relating cosmological concepts to their own lives.

Eddington, Arthur. *The Expanding Universe*. Reprint. Cambridge, England: Cambridge University Press, 1987. Reissue of a classic early work supporting the theory of an expanding universe by a foremost scientist of the time. Contains a foreword by William McCrea, placing the work in its historical perspective.

Ferreira, Pedro G. *The State of the Universe: A Primer in Modern Cosmology*. London: Phoenix, 2007. Oxford lecturer Ferreira presents a history of cosmology, examining the complexities that concepts such as dark matter and dark energy have imposed on a once "simple" Einsteinian universe ruled by relativity.

Gasperini, Maurizio. *The Universe Before the Big Bang: Cosmology and String Theory*. Berlin: Springer, 2008. A fascinating exploration of what might have happened prior to the explosion that formed the universe, which looks to string theory and other modern mathematical models to postulate that the universe was not born with the big bang but rather was well advanced in its overall evolution. Presented with nontechnical language for nonspecialists.

Gribbin, John. *In Search of the Big Bang*. New York: Bantam Books, 1986. Explains the search for an understanding of the nature of the universe. "Part One: Einstein's Universe" gives a historical understanding of events leading to the discovery of the expanding universe. For the general reader.

Hawking, Stephen W. *A Brief History of Time: From the Big Bang to Black Holes*. New York: Bantam Books, 1988. Geared for the general reader by one of the leading theoretical physicists of the twentieth century. Provides a nonmathematical step-by-step explanation of the expanding universe and why the universe must expand.

Lemoine, M., J. Martin, and P. Peter, eds. *Inflationary Cosmology*. New York: Springer, 2008. This collection of papers, by both venerable and younger astrophysicists and cosmologists, focuses on that period during which the early universe expanded at a greatly accelerated rate before settling down to a slower rate of expansion. Presents several different scenarios.

Liddle, Andrew, and Jon Loveday. *The Oxford Companion to Cosmology*. New York: Oxford University Press, 2008. An indispensable A-Z reference, consisting of more than 350 entries from antimatter to WIMPs, as well as individual physicists and other scientists who have advanced the field. Heavily illustrated with almost two hundred halftones and diagrams; includes cross-references and Web links.

North, John. *Cosmos: An Illustrated History of Astronomy and Cosmology*. Chicago: University of Chicago Press, 2008. Emphasizes the astrophysics of cosmology, with a focus on physical and mathematical concepts that are key to understanding classical field theory. Also describes experimental techniques and results. Technical.

Schneider, Peter. *Extragalactic Astronomy and Cosmology: An Introduction*. New York: Springer, 2006. A textbook focusing on galaxies, clusters, and superclusters, beginning with the Milky Way. Covers their evolution, formation, and distribution, supported by beautiful color illustrations.

Weinberg, Steven. *Cosmology*. New York: Oxford University Press, 2008. The Nobel physics laureate presents the subject in two parts: one covering the isotropic and homogeneous "average" universe; the second, departures from the average universe. Provides detailed coverage of recombination, microwave background polarization, leptogenesis, gravitational lensing, structure formation, and multifield inflation. Includes mathematical calculations. Appendixes review general relativity, the Boltzmann equation, and sample problems. For advanced students and professionals.

Wudka, Jose. *Space-Time, Relativity, and Cosmology*. New York: Cambridge University Press, 2006. A history of relativistic cosmology from ancient times to Einstein. The nonmathematical approach emphasizes concepts over calculations yet explain the ideas clearly, applying them to research topics in cosmology. Designed for students from high school through college.

See also: Big Bang; Cosmic Rays; Cosmology; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; General Relativity; Interstellar Clouds and the Interstellar Medium; Milky Way; Novae, Bursters, and X-Ray Sources; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Universe: Evolution; Universe: Structure.

Universe: Structure

Category: The Cosmological Context

Galaxies are gravitationally bound assemblages of millions to tens of trillions of stars. Galaxies are not scattered randomly across the universe but instead are grouped together in galaxy clusters, consisting of several tens to many thousands of individual galaxies. There is growing evidence of connective patterns between galaxy clusters stretching across regions at least as large as many hundreds of millions of light-years. Explaining the observed features of this large-scale structure is a challenge to cosmologists that puts limits on acceptable models of the origin and evolution of the universe.

OVERVIEW

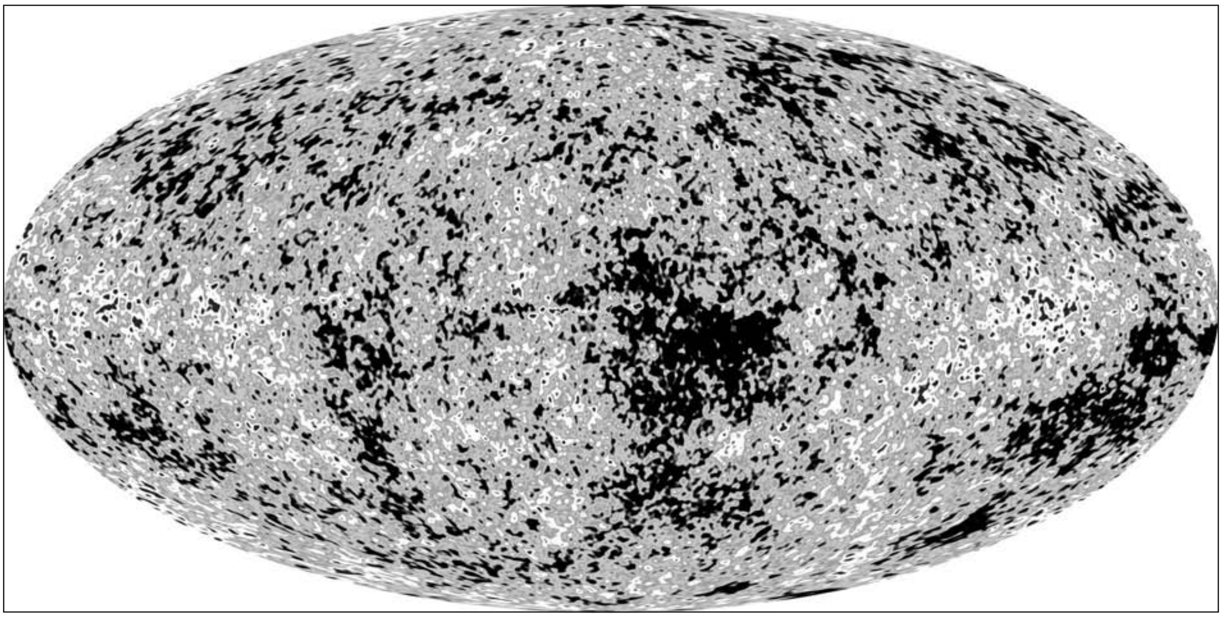
The contents of the universe are arranged in a hierarchy of structures. Stars, of which our Sun is a familiar example, are hot balls of gas that generate energy through nuclear fusion reactions. There is growing evidence that many stars are orbited by families of planets, analogous to our solar system. Stars (and their families of planets, if they have them) are gravitationally bound together into galaxies, vast collections of millions to tens of trillions of stars, spanning thousands to hundreds of thousands of light-years. Our Sun and solar system are part of the Milky Way galaxy, a moderately large spiral galaxy consisting of several hundred billion stars along with gas and dust.

Galaxies in turn are gravitationally bound together into galaxy clusters, containing from several tens to many thousands of galaxies. The Milky Way galaxy is part of the Local Group, a small galaxy cluster containing about forty members. Galaxy clusters group together to form superclusters. The Local Group is part of the Virgo supercluster, containing several tens of thousands of individual galaxies and spanning more than 100 million light-years. On a still larger scale, galaxies, clusters, and superclusters appear to be arranged in a network of filaments and walls surrounding large, nearly empty voids or bubbles hundreds of millions of light-years across. This overall large-scale structure has been described as "frothy," or analogous to Swiss cheese with all its holes.

Understanding the structure and evolution of the universe requires evidence about the distribution of matter on the largest scales. Most models of the universe satisfy the cosmological principle. This is the assumption that, at any moment, the large-scale structure of the universe looks the same from all places (homogeneity) and in all directions (isotropy).

When Albert Einstein published his general theory of relativity in 1916, in which gravity is treated as a warping of space-time, almost immediately it was recognized that it would have a profound impact on cosmological models. Einstein himself tried applying it to a static universe and found that he needed to add an arbitrary term, which was called the cosmological term or cosmological constant, to obtain a static solution.

In 1922, the Russian mathematical physicist Alexander Alexandrovich Friedmann and in 1927 the Belgian priest and cosmologist Georges Lemaître independently derived two classes of homogeneous and isotropic solutions (without the cosmological constant) in which the universe expands. In one of these classes, space has a uniform positive curvature and finite extent though no boundary; the universe is said to be closed because it expands to some maximum size and then contracts. The two-dimensional surface (not including the interior volume) of a three-dimensional sphere is such a space; motion on the surface is never blocked by a barrier (perimeter line), but the area of the surface is finite. In the other class, space is negatively curved and of infinite extent; the universe is said to be open because it expands forever. In two dimensions, such a surface is termed "hyperboloid"; a saddle is an example of a finite part of such a surface which mathematically extends to infinity if it lacks a boundary curve. In 1932, Einstein and the Dutch astronomer Williem de Sitter proposed a third type of expanding model without the cosmological constant. In it, space is flat and infinite, geometry is Euclidean, and the universe just barely expands forever; it is the boundary case between the two Friedmann-Lemaître solutions. In two



In 2003, the Wilkinson Microwave Anisotropy Probe produced this high-resolution map of the universe's microwave radiation only 380,000 years after the big bang. These data helped to prove that the age of the universe is 13.7 billion years; that 73 percent of the universe is composed of dark energy, 23 percent of cold dark matter, and 4 percent of atoms; and that the universe will continue to expand. (NASA/WMAP Science Team) dimensions, the surface of a plane is flat and infinite. These are the only three possible spatial geometries and types of models that can be solutions of Einstein's gravitational field equations under the assumptions of homogeneity and isotropy.

The common feature of these models is that their scale factors (the distance between representative points, such as clusters of galaxies) change with time; the universe must be either expanding or contracting. In 1929, Edwin Powell Hubble published the first observational evidence that the universe is indeed expanding. Using the 100-inch reflecting telescope on Mount Wilson (then the largest telescope in the world), he found that the more distant a galaxy is, the greater its spectrum is redshifted. This means all features in the spectrum are shifted to longer than normal wavelengths, and this can be due to motion away from us.

If the universe is expanding, the scale factor was smaller in the past, and the local densities of both matter and radiation were greater. However, the mass density of matter varies inversely as the scale factor cubed, while the energy density of radiation varies inversely as the scale factor to the fourth power. At sufficiently early times when the scale factor was smaller, the density of radiation was greater than the density of matter, and the universe was dominated by radiation. With time, as the scale factor increased and the universe expanded, the density of radiation decreased faster than the density of matter did, and eventually the universe became dominated by matter, as it is today.

The universe today is "lumpy" on a variety of scales: stars, galaxies, galaxy clusters, superclusters, and "walls" and "voids." In recent years, there has been growing interest in trying to understand how an initially homogeneous big bang could produce the lumpy inhomogeneities observed today. This is called the homogeneity problem.

APPLICATIONS

Galaxy surveys carried out to great distances show that the density of galaxies in voids is typically a factor of ten less than average, and the density in the narrow but long walls and filaments is typically a factor of ten thousand greater than average. Explaining the origin of such variations in density from an early universe that was amazingly homogeneous is a major challenge in modern cosmology.

The cosmic microwave background radiation (called the CMB or CBR for short), coming from a time a few hundred thousand years after the big bang when the matter in the expanding universe became transparent, is remarkably uniform and isotropic. However, detailed observations of it made by the Cosmic Background Explorer (COBE) satellite, the Wilkinson Microwave Anisotropy Probe (WMAP) satellite, and the Balloon Observations of Millimetric Extragalactic Radiation and Geomagnetics (BOOMERANG) project in Antarctica do reveal small variations. The data perfectly fit a blackbody Planck curve for a temperature of 2.735 kelvins. Removing a slight asymmetry in temperature (about 0.007 kelvins) in opposite directions, presumably due to the motion of our solar system through the background radiation, leaves very small temperature fluctuations on the order of 10 micro-kelvins, about one degree in angular size. The hotter regions had slightly higher densities, and they are about the right size to develop into clusters and superclusters of galaxies.

The initial formation of the hotter, denser regions and their evolution into galaxy clusters and superclusters may involve dark matter. This is the name now used to refer to what formerly was called "missing mass." In many situations in astronomy, the amount of matter that can be detected through the electromagnetic radiation that it emits (whether radio waves, visible light, or other wavelengths) is much less—typically by a factor of about 5 to 50—than what is needed gravitationally to hold galaxies, clusters of galaxies, and superclusters together and to account for gravitational lensing of distant objects. Most astrophysicists consider the gravitational estimates reasonably well established, and thus believe the mass is not "missing" but simply is not emitting electromagnetic radiation. The challenge is to explain the nature of the dark matter, which is much more abundant than ordinary luminous matter.

Various observational tests suggest that nonluminous ordinary matter (perhaps in the form of boulders, planet-sized objects, black dwarfs, black holes, and other known nonemitting entities) can account for no more than about 10 to 12 percent of dark matter. Thus most dark matter must be in some more exotic form. One suggestion involves weakly interacting massive particles (WIMPs), subatomic particles that, like neutrinos, interact very rarely with ordinary matter (other than gravitationally) but are much more massive than neutrinos.

Whatever dark matter is, it seems to be necessary to give us the universe we observe today. Temperature (and hence density) fluctuations revealed in the background radiation are about the right size to develop into galaxy clusters and superclusters, but they do not contain enough ordinary matter to contract gravitationally into protogalaxies as quickly as galaxies seem to have formed after the big bang. Adding a lot of dark matter to these density fluctuations stimulates the rapid growth of galaxies because of the increased gravitational attraction.

Temperature and density fluctuations recorded in the background radiation probably are due to very small quantum fluctuations that occurred spontaneously during the first 10^{-35} second after the big bang, and then were magnified by many orders of magnitude during the era of cosmic inflation, when the universe expanded exponentially in the next 10^{-30} second. Since ordinary matter was opaque to electromagnetic radiation until a few hundred thousand years after the big bang, it would have been buffeted and kept smoothed out by the strong radiation field until finally the ordinary matter became transparent and decoupled from the radiation. However, since dark matter does not seem to interact with electromagnetic radiation, it would have been able to collect in the magnified quantum fluctuations as soon as inflation ended and thus built up density concentrations to attract ordinary matter later.

Another idea that may be relevant to the formation of large-scale structure in the universe is cosmic string theory. According to string theory, space-time has ten or eleven dimensions, but most are "rolled up" or compacted so that only the familiar four space-time dimensions (length, height, width, and time) are noticeable. Cosmic strings are hypothetical long, thin, linelike concentrations of unbroken symmetry left over from the spontaneous symmetry breaking that occurred when the electromagnetic, weak nuclear, and strong nuclear forces separated to bring the grand unification era to a close. Specific multidimensional modes of vibration of cosmic strings are thought to be manifested as all the particles and forces in the universe. In the early universe, cosmic strings may have served as "seeds" for the formation of long concentrations of matter, like droplets condensing on a wire, which later evolved into filamentary chains of galaxies.

CONTEXT

The history of the quest to understand the structure of the universe has been a progression toward recognition of ever more subtle organization at ever-larger scales. Just about all cultures and societies have divided naked-eye stars of the night sky into patterns we call constellations. In early Greco-Roman cosmology (and in many others), stars were attached to the inside of a hollow sphere—the celestial sphere—that enclosed the Earth fixed at the sphere's center. In the early 1600's, Galileo, in the first recorded use of telescopes to systematically study the sky, discovered there were many stars too faint to be seen with the unaided eye. Gradually during the 1600's, the idea developed that stars were similar to our Sun, and therefore their different apparent brightnesses meant that they were at different distances from us. Consequently the stars in a given pattern or constellation might not be a real grouping in space but could be at very different distances from us.

However, this idea could not be confirmed until the distances of stars could be measured directly. The first successful measurements of stellar distances were made in 1838 and 1839, independently by Friedrich Wilhelm Bessel in Germany, Thomas Henderson in South Africa, and Friedrich Georg Wilhelm von Struve in Russia. They all used the method of trigonometric parallax. That involved measuring small changes in the apparent positions of nearby stars relative to more distant stars as seen from

Leavitt, Shapley, and the Period-Luminosity Scale

In 1902, Henrietta Swan Leavitt became a permanent staff member at Harvard College Observatory. She studied variable stars, stars that change their luminosity (brightness) in a fairly predictable pattern over time. During her tenure at Harvard, Leavitt observed and photographed nearly 2,500 variable stars, measuring their luminosities over time. She was equipped with photographs of the Large and Small Magellanic Clouds collected from Harvard's Peruvian observatory. The Magellanic Clouds are very small galaxies visible in the Southern Hemisphere and close to the Milky Way. The Small Magellanic Cloud contained seventeen Cepheid variables having very predictable periods ranging from 1.25 days to 127 days. Leavitt carefully measured the brightening and dimming of the seventeen Cepheids during their respective periods. She collected photographs of other Cepheids in the Magellanic Clouds and made additional period-luminosity studies. In a circular dated March 3, 1912, she stated:

> The measurement and discussion of these objects present problems of unusual difficulty, on account of the large area covered by the two regions, the extremely crowded distribution of the stars contained in them, the faintness of the variables, and the shortness of their periods. As many of them never become brighter than the fifteenth magnitude, while very few exceed the thirteenth magnitude at maximum, long exposures are necessary, and the number of available photographs is small. The determination of absolute magnitudes for widely separated sequences of comparison stars of this degree of faintness may not be satisfactorily completed for some time to come. With the adoption of an absolute scale of magnitudes for stars in the North Polar Sequence, however, the way is open for such a determination.

Ejnar Hertzsprung of the Leiden University in the Netherlands and Henry Norris Russell of the Mount Wilson Observatory in Pasadena, California, had independently discovered a relationship between a star's luminosity and its spectral class (that is, color and temperature). Together, their experimental results produced the Hertzsprung-Russell diagram of stellar luminosities, the astronomical equivalent of chemistry's periodic table. According to their classification scheme, most stars lie along the "main sequence," which ranges from extremely bright blue stars ten thousand times brighter than the Sun to very dim red stars one hundred times dimmer than the Sun. Cepheid variables fell toward the cooler, red end of the main sequence.

Leavitt carefully measured the luminosities and cyclic periods of changing luminosity for each of many Cepheid variables from the Magellanic Clouds. From her careful measurements, she graphically plotted Cepheid luminosity against Cepheid period. She noticed "a remarkable relation between the brightness of these variables and the length of their periods. . . . the brighter variables have the longer periods." She had discovered that a Cepheid's apparent luminosity is directly proportional to the length of its period, or the time it takes to complete one cycle of brightening and dimming.

Harlow Shapley, an astronomer at the Mount Wilson Observatory, measured the distances of moving star clusters containing Cepheids, then related the Cepheid distances to Cepheid period-luminosity data. From these experiments, Shapley constructed a Cepheid period-absolute luminosity curve, which made it possible to plot a Cepheid variable having a specific measured period and obtain its absolute luminosity. Knowing the Cepheid's apparent and absolute luminosities, one can instantly calculate its distance and, therefore, the distances of all the stars in the star cluster containing that particular Cepheid variable.

The distances to Cepheid variables in the Milky Way and other galaxies were soon determined. Shapley used Cepheid distances to demonstrate that the center of the Milky Way is directed toward the constellation Sagittarius and that the Sun is located approximately thirty thousand light-years from the galactic center. Edwin Powell Hubble applied the technique to obtain estimates of the distances between our galaxy and others, which led to his monumental astronomical discovery that the universe is expanding.

Earth at various points in its orbit around the Sun. Trigonometric parallax is still the most assumption-free method of measuring distances, but it is limited by the ability to measure small angular shifts accurately. Hipparcos, the High-Precision Parallax Collecting Satellite operated by the European Space Agency, has been able to measure parallaxes with reasonable accuracy out to distances of about 1,600 light-years. Beyond that distance, other methods must be used, but their calibration ultimately is tied in to the distances obtained by trigonometric parallaxes. Thus constellations came to be seen as merely convenient direction indicators from our vantage point on Earth, not physical associations of stars.

In the early 1600's, Galileo discovered telescopically that the hazy white band of light known as the Milky Way actually consisted of lots of faint stars. Because the Milky Way forms a great circle band of light around the sky, by the 1700's its overall shape was described using terms like sheet, disk, millwheel, and grindstone. In 1784, William Herschel tried to determine its size and shape by counting the stars seen in various directions. These early models all placed the Sun near the Milky Way's center. It was not until 1918 that Harlow Shapley found that the Milky Way galaxy was much larger than previous estimates and the Sun was located far from the center.

At this time, there still was uncertainty about whether the Milky Way galaxy comprised the entire universe or not. As far back as the 1700's, there had been speculation, by Immanuel Kant and others, that the spiral nebulae might be other large collections of stars like the Milky Way. During the 1920's, Edwin Hubble, using the Mount Wilson 100-inch reflecting telescope, was able to resolve individual stars in some of the spiral nebulae; moreover, some of the stars were Cepheid variable stars, which could be used to determine distances. The spiral nebulae were shown to be located far beyond the Milky Way galaxy and hence were galaxies comparable to the Milky Way. A few years later, Hubble found that most galaxy spectra were redshifted, thus confirming the cosmological models of an expanding universe derived from Einstein's general relativity.

Ever since, central goals of observational and theoretical cosmology have been to produce accurate maps of the distribution of these galaxies and explanations of the origin of the features in this distribution. Pioneering work by George Abell and others, beginning in the 1950's, showed that galaxies were grouped into clusters, some rich with thousands of member galaxies and others poor with fewer than one hundred members. Some clusters are regular, with a spherical shape and many galaxies concentrated near the center; others are irregular, with galaxies scattered across an extended region of space. In the 1980's, Margaret Geller and John Huchra at Harvard-Smithsonian Center for Astrophysics began mapping the large-scale distribution of galaxies out to great distances. They, and teams at other institutions, have shown that out to distances of several billion light-years, superclusters of galaxy clusters are arranged in filaments and walls surrounding large, nearly empty voids.

Understanding the origin of structure in the universe on such large scales challenges the creativity of cosmologists. Although the exact details are not yet clear, it seems that all the patterns of organization seen in the universe today were determined by events and processes in its earliest moments after the big bang.

John J. Dykla and Richard R. Erickson

FURTHER READING

Belusevic, Radoje. *Relativity, Astrophysics, and Cosmology*. Weinheim, Germany: Wiley-VCH, 2008. Addresses the interrelationship of the now entwined disciplines of astrophysics, particle physics, cosmology, and relativity theory, which have combined to form advanced theories of the origin and evolution of the universe.

Bennett, Jeffrey, et al. *The Cosmic Perspective*. 5th ed. San Francisco: Pearson/Addison-Wesley, 2008. Structured around a set of two-page figures dubbed "cosmic contexts," this interactive resource uses zoom-in illustrations to orient students to various images of the universe in relation to one another, while covering all the basics of cosmology. For general and introductory audiences. Chapters include "Light and Matter: Reading Messages from the Cosmos," "A Universe of Galaxies," "Dark Matter," "Dark Energy," and "The Fate of the Universe."

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Several chapters address the development of structure in the universe.

Cohen, Nathan. *Gravity's Lens: Views of the New Cosmology*. New York: John Wiley & Sons, 1988. A book for the general reader by a researcher in general relativity and cosmology, this clear and well-illustrated volume features extensive discussion of the evidence for large-scale structure, and prospects for refined and more extensive observations in the future.

Drexler, Jerome. *Discovering Postmodern Cosmology: Discoveries in Dark Matter, Cosmic Web, Big Bang, Inflation, Cosmic Rays, Dark Energy, Accelerating Cosmos*. Boca Raton, Fla.: Universal, 2008. Cosmologist Drexler proposes a plausible and unique correlation of the seven mysterious areas of cosmological research listed in the book's subtitle. For all open-minded audiences.

Duncan, Todd, and Craig Tyler. *Your Cosmic Context: An Introduction to Modern Cosmology*. San Francisco: Pearson Addison Wesley, 2009. An introductory texbook for studies in modern cosmology that engages students by relating cosmological concepts to their own lives.

Ferreira, Pedro G. *The State of the Universe: A Primer in Modern Cosmology*. London: Phoenix, 2007. Oxford lecturer Ferreira presents a history of cosmology, examining the complexities that concepts such as dark matter and dark energy have imposed on a once "simple" Einsteinian universe ruled by relativity.

Gasperini, Maurizio. *The Universe Before the Big Bang: Cosmology and String Theory*. Berlin: Springer, 2008. A fascinating exploration of what might have happened prior to the explosion that formed the universe, which looks to string theory and other modern mathematical models to postulate that the universe was not born with the big bang but rather was well advanced in its overall evolution. Presented with nontechnical language for nonspecialists.

Lemoine, M., J. Martin, and P. Peter, eds. *Inflationary Cosmology*. New York: Springer, 2008. This collection of papers, by both venerable and younger astrophysicists and cosmologists, focuses on that period during which the early universe expanded at a greatly accelerated rate before settling down to a slower rate of expansion. Presents several different scenarios.

Liddle, Andrew, and Jon Loveday. *The Oxford Companion to Cosmology*. New York: Oxford University Press, 2008. An indispensable A-Z reference, consisting of more than 350 entries from antimatter to WIMPs, as well as individual physicists and other scientists who have advanced the field. Heavily illustrated with almost two hundred halftones and diagrams; includes cross-references and Web links.

North, John. *Cosmos: An Illustrated History of Astronomy and Cosmology*. Chicago: University of Chicago Press, 2008. Emphasizes the astrophysics of cosmology, with a focus on physical and mathematical concepts that are key to understanding classical field theory. Also describes experimental techniques and results. Technical.

Schneider, Peter. *Extragalactic Astronomy and Cosmology: An Introduction*. New York: Springer, 2006. A textbook focusing on galaxies, clusters, and superclusters, beginning with the Milky Way. Covers their evolution, formation, and distribution, supported by beautiful color illustrations.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. An introductory astronomy textbook. Divided into short units on specific topics, it offers extensive coverage throughout on the development of structure in the universe.

Weinberg, Steven. *Cosmology*. New York: Oxford University Press, 2008. The Nobel physics laureate presents the subject in two parts: one covering the isotropic and homogeneous "average" universe; the second, departures from the average universe. Provides detailed coverage of recombination, microwave background polarization, leptogenesis, gravitational lensing, structure formation, and multifield inflation. Includes mathematical calculations. Appendixes review general relativity, the Boltzmann equation, and sample problems. For advanced students and professionals.

Wudka, Jose. *Space-Time, Relativity, and Cosmology*. New York: Cambridge University Press, 2006. A history of relativistic cosmology from ancient times to Einstein. The nonmathematical approach emphasizes concepts over calculations yet explain the ideas clearly, applying them to research topics in cosmology. Designed for students from high school through college.

See also: Big Bang; Cosmic Rays; Cosmology; Electromagnetic Radiation: Nonthermal Emissions; Electromagnetic Radiation: Thermal Emissions; General Relativity; Interstellar Clouds and the Interstellar Medium; Milky Way; Novae, Bursters, and X-Ray Sources; Space-Time: Distortion by Gravity; Space-Time: Mathematical Models; Universe: Evolution; Universe: Expansion.

Uranus's Atmosphere

Categories: Planets and Planetology; The Uranian System

Uranus is the seventh planet from the Sun. It shares much in common with Jupiter and Saturn, but it is also significantly different from the larger Jovian, or "gas giant," planets. Its atmosphere is composed mainly of hydrogen and helium, but its color is governed by selective absorption of light by methane, which is abundant in greater measure in Uranus's atmosphere than in either Jupiter's or Saturn's.

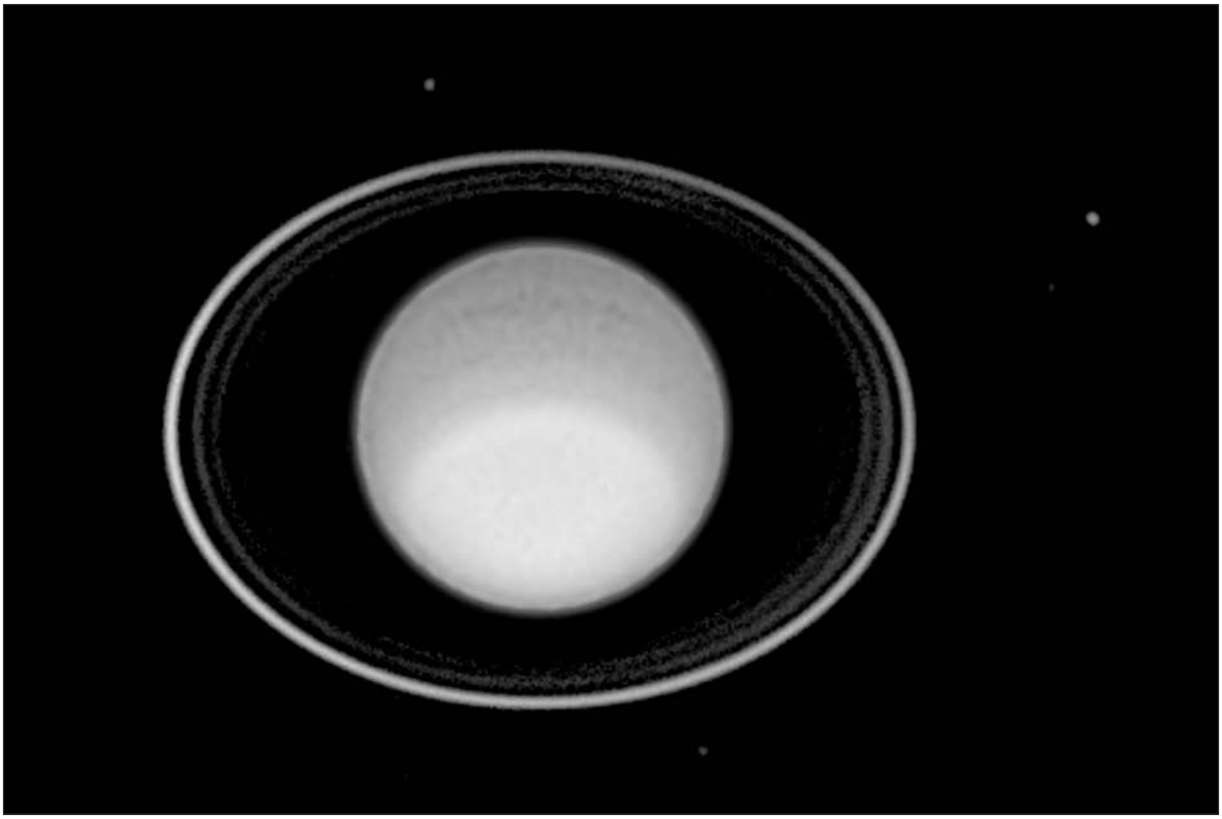
OVERVIEW

The planet Uranus was the first to be discovered with a telescope. Its existence was declared by Sir William Herschel on March 13, 1781. After several proposed names, the most curious of which was a proposed reference to the King of England, George III, the planet was named Uranus. From mythology Uranus is the father of Saturn and grandfather of Jupiter.

Uranus is the third largest planet in the solar system. With an orbit that varies from 18.4 to 20 astronomical units (AU, or the mean distance from the Earth to the Sun, namely 150 million kilometers), it takes Uranus eighty-four years to complete one revolution about the Sun. Naturally, at this greater distance from the Sun, Uranus receives far less solar radiation than Jupiter and Saturn. Nevertheless, its location suggested to early researchers that the composition and nature of Uranus were similar to those of Jupiter and Saturn.

Superficially, that is true. Uranus is composed largely of hydrogen and helium. However, the atmosphere of Uranus has been determined to be colder than that of Jupiter and Saturn and has a less dynamic structure than the turbulent atmosphere of Jupiter or the pastel banding of Saturn. Uranus's atmospheric temperature can drop to 49 kelvins. In addition to the preponderance of hydrogen and helium, the atmosphere has a larger amount of ices and hydrocarbon than do the atmospheres of Jupiter and Saturn. Ices include water, ammonia, ammonium hydrosulfide, and methane. Selective absorption of radiation, in good measure by methane, results in the planet's pale bluish-green appearance.

By compositional abundance, Uranus is 83 percent hydrogen, 15 percent helium, 2.3 percent methane. The total also includes other, low-concentration gases and hydrocarbons and does not add up precisely to 100 percent given significant uncertainties about the abundance of hydrogen and helium. Hydrocarbons that appear only in trace amounts include ethane, acetylene, methyl acetylene, and diacetylene. These and other hydrocarbons are thought to be produced in the upper atmosphere by photolysis of methane under incident solar ultraviolet light. Carbon monoxide and carbon dioxide have also been detected. Like the planet's water vapor, carbon dioxide and carbon monoxide must have been acquired by impacting comets and infalling dust. All totaled, Uranus has a carbon content, primarily found in the atmosphere,



In 1995, the Hubble Space Telescope revealed much about Uranus's atmosphere: Infrared images showed that it is composed of hydrogen and traces of methane; the inner atmosphere is clear, an intermediate yellow layer is hazy, and a very thin outer layer is red. In addition, Uranus's rings are bright in the infrared. (NASA/Erich Karkoschka, University of Arizona)

somewhere between twenty and thirty times that of solar abundance.

One thing that makes Uranus particularly curious is the fact that its rotational axis is tilted 97.8° from the perpendicular to the ecliptic plane. This tilt is why many refer to Uranus as the planet that rotates on its side. There is no universally accepted explanation for this high degree of tilt, but many believe the planet was knocked on its side by a collision with a large body early in the Uranian system's development. This curious tilt means that for roughly half of each orbit the north pole receives solar radiation, and for roughly half of the rest of the orbit the south pole is in sunlight. This makes for unusual seasons and atmospheric dynamics. Although the planet's interior rotates once every 17 hours 14 minutes, the atmosphere rotates differentially. Features in the upper atmosphere have been clocked at as much as 0.25 kilometer per second and thus may experience a full rotation in less than 14 hours.

KNOWLEDGE GAINED

Earth-based telescopic studies of Uranus revealed it to have a bizarre orientation of its rotational axis, several relatively small satellites, and an orbital period of 84.3 years. In 1977, observations made from an aircraft-based telescope as Uranus occulted a star revealed the presence of dark rings around the mysterious seventh planet from the Sun. That same year the Voyager 2 spacecraft launched on an approved mission to fly by both Jupiter and Saturn. The National Aeronautics and Space Administration (NASA) originally proposed sending an armada of sophisticated spacecraft on what had been termed the "Grand Tour."

This "tour" referred to the fact that every 176 years, planetary alignments are such that gravitational slingshot maneuvers in the outer solar system can be used to send spacecraft to investigate all the outer planets from Jupiter to Pluto. Unfortunately, that ambitious plan was not funded, but NASA was given authorization to build two modest Voyager spacecraft for exhaustive investigations of Jupiter and Saturn. When Voyager 1 was successful at both gas giants, the Voyager 2 spacecraft was targeted through the Saturn system in such a way as to make possible a flyby of Uranus and Neptune.

Atmospheric structure is often discussed in terms of either pressure levels or temperature or both. If one defines Uranus's "surface" as the site where the pressure is 1 bar (1 Earth atmosphere, or 10^5 pascals), then that atmosphere can be described as follows. Uranus has a troposphere found from –300 to 50 kilometers above the surface, where the pressure varies from 100 to 0.1 bar, respectively. A stratosphere exists between 50 and 400 kilometers, where the pressure varies from 0.1 to 10^{-10} bar. Then, from 400 kilometers out to as much as two planet radii, or roughly 50,000 kilometers, is the thermosphere and corona, where the pressure dwindles down to near vacuum from the upper stratospheric level of 10^{-10} bar.

One might think that because its poles receive more solar illumination than the equatorial region, Uranus would be warmer at the pole presently facing toward the Sun, but that is not the case. Near the equator is the planet's only portion to experience fairly rapid day-night variation due to the excessive tilt of Uranus. Near the equator the warmest temperatures are recorded. Upper atmospheric temperatures near the equator can rise to 57 kelvins. Why the equatorial region is warmer than the illuminated polar region is currently unknown.

The Hubble Space Telescope routinely was used by planetary scientists to examine Uranus for features and changes in those features within the planet's atmosphere. In 1998 an image credited to NASA and Erich Karkoschka of the University of Arizona revealed on the order of twenty clouds in Uranus's atmosphere. That was rather remarkable; prior to that time, in the entire history of Uranus observations, there had been fewer than that number of clouds seen in the planet's usually unremarkable-looking atmosphere. This Hubble image was taken in infrared and clearly showed the planet's rings and many of its known satellites. In addition to the clouds, it revealed a bright band circling the planet. Wind speeds of clouds near the band were determined to be in excess of 500 kilometers per hour. One of the clouds seen in this infrared image was the brightest Uranian cloud ever observed.

In 2006, Hubble images, in concert with near-infrared observations made using the ground-based Keck telescope, revealed a new, more dynamic picture of Uranus. The planet was seen to have a great dark spot and some degree of banding. These observations were summarized by astronomer Heidi Hammel at a Hubble science overview briefing held in 2002, about a month in advance of what was then proposed to be the final shuttle servicing mission to the Hubble. Hammel used her appearance at that briefing as an opportunity to stress how much Hubble had already done to change the picture of an inactive Uranian atmosphere as presented by Voyager 2 in 1986. Hammel and other astronomers had detected an increase in the number and scope of clouds in the ice giant's atmosphere, in addition to finding that great dark spot. She expressed enthusiasm about extended Hubble operations advancing understanding of Uranus, perhaps the planet about which the least is known. This increased activity appears related to seasonal changes as Uranus orbits the Sun.

CONTEXT

Uranus was the first planet for which clear records for discovery exist. Although Herschel was not the first to note Uranus in astronomical records, he was the first to identify it correctly as a planet and not a comet or unidentified star, as others had done previously. From the time of its discovery to the dawn of the space age, little could be learned about Uranus from ground-based telescopes. However, observations of Uranus's orbit about the Sun led to the recognition that there was good reason to believe that it was not the last planet to be discovered in the solar system. Based on gravitational perturbations in the orbit of Uranus, Neptune was discovered by and large by mathematical analysis. Observations verified the correctness of those calculations. Uranus and Neptune were believed to be very similar. Both displayed a bluish-green tint in telescopic views. Spectroscopic analysis indicated that both planets had atmospheres different from those of Jupiter and Saturn. Like their larger gas giant cousins, Uranus and Neptune were known to have atmospheres rich in hydrogen and helium, but their bluish-green color was identified as due to extensive absorption of red light by methane.

Voyager 2 provided the greatest portion of current understanding about the Uranian system. Planetary scientists interested in Uranus await a return mission, most likely an orbiter, perhaps with a lander probe for one of the icy satellites and an atmospheric probe to ram through the upper atmosphere of Uranus and conduct measurements until it is crushed. Ground-based observations and imaging by the Hubble Space Telescope continue in the meantime. The biggest question about detection and observation of clouds in Uranus's atmosphere centers on the source of energy driving those storms, since Uranus's internal heat flow appears to be insufficient to cause such airflow.

David G. Fisher

FURTHER READING

Elkins-Tanton, Linda T. *Uranus, Neptune, Pluto, and the Outer Solar System*. New York: Chelsea House, 2006. This book explores the Sun's relationship with the three outer planets and their moons, considering these planets as recorders of the formation of the solar system. Aimed at a general or high school audience. Illustrations, bibliography, index.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that had investigated all planets by the publication date. Takes an astrophysical approach to place our solar system in a broad context as just one member of similar systems throughout the universe.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college text on astronomy, somewhat more advanced than many introductory texts but with a wealth of detail and excellent diagrams. Chapters 6 through 16 describe the solar system, including Uranus and what is know about its atmosphere. Comes with a CD-ROM.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Focuses on Jupiter, Saturn, Uranus, and Neptune and their atmospheres. Suitable as a textbook for upper level college courses in planetary science. Filled with figures and photographs. Available to the serious general audience.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar-system astronomy suitable for an introductory college course. Accessible to nonspecialists as well. Filled with supplemental learning aids and solved student exercises.

Miller, Ron. *Uranus and Neptune*. Brookfield, Conn.: Twenty-First Century Books, 2003. Considers Uranus and its satellites in comparison with other gas giants, especially Neptune, including their atmospheres.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. Geared for the undergraduate college student. Planetary atmospheres are treated as important physical features of the various members of the Sun's family. They are discussed individually in the context of what is known about each planet's characteristics and with regard to theories about their evolution and the evolution of the entire solar system. Comprehensive for the average reader.

Tocci, Salvadore. *A Look at Uranus*. New York: Franklin Watts, 2003. As part of the Out of This World series, this book covers all aspects of the planet Uranus from its discovery through 2002. Includes several photographs. Suitable for all readers.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Eclipses; Infrared Astron-

omy; Jovian Planets; Planetary Atmospheres; Planetary Formation; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Telescopes: Ground-Based; Uranus's Interior; Uranus's Magnetic Field; Uranus's Rings; Uranus's Satellites; Uranus's Tilt.

Uranus's Interior

Categories: Planets and Planetology; The Uranian System

Uranus, the seventh planet from the Sun, has much in common with the other Jovian gas giants, but its interior is different from the interiors of the larger planets Jupiter and Saturn and more like that of Neptune.

OVERVIEW

Uranus is largely composed of hydrogen and helium and is considered to be a Jovian planet, or gas giant. The interior of Uranus differs significantly from that of Jupiter and Saturn, however. It shares much more in common with Neptune. Like the interior of Jupiter and Saturn, Uranus's interior cannot be directly sampled. Models of the interiors of the Jovian planets are inferred by external observation. For example, a magnetic field tells much about the nature and physical characteristics of a given planet's interior.

Accurate determinations of orbital motions of Uranus's satellites led to a precise value of the planet's mass. With its size slightly in excess of Neptune, Uranus is the third largest planet in the solar system, but the second least dense. Uranus has a mass 14.5 times that of Earth. With a mean radius four times the Earth's, its overall density is 1.27 grams per cubic centimeter. Of course, the different structures of Uranus, from the core to the upper atmosphere, have specific characteristics of their own. Nevertheless, in gross terms, only Saturn is less dense than Uranus as a planet. Uranus is composed, like Jupiter and Saturn, primarily of hydrogen and helium, although the percentages of both are different in Uranus from the percentages of Jupiter and of Saturn. Uranus contains methane, water vapor, ammonia, carbon monoxide, and carbon dioxide. Hydrocarbons such as ethane, acetylene, methyl acetylene, and diacetylene are also present, presumably created by photolysis of methane in the atmosphere under illumination of solar ultraviolet light.

Like Jupiter and Saturn, hydrogen in Uranus is found in gaseous form in the atmosphere, but deep inside the planet at greater pressure hydrogen may exist in more exotic forms. Whether helium is found in forms other than as a simple gas is a matter of debate. Because of their size, temperature, pressure, compositional, and interior structural differences from Jupiter and Saturn, Uranus and Neptune both are often referred to as ice giants.

Several models exist for Uranus's interior. They agree on major features and differ in only minor ways. A great deal of the interior is composed of water, ammonia, and methane, with the percentage of water included in each model varying. That water percentage ranges between 9.3 to 13.5 Earth masses, depending on which model a scientist supports. With only 0.5 to 1.5 Earth masses of hydrogen and helium in the atmosphere and interior, that leaves between 0.5 to 3.7 Earth masses for what is considered rocky material.

Under the atmosphere of hydrogen, helium, and methane are two inner layers. First, an icy mantle actually incorporates the majority of the planet's mass. Under that is the rocky core. The core represents less than 20 percent of the planet's radius, and the relatively thin upper atmosphere of gases represents another 20 percent of the planet's radius. This leaves the mantle as 60 percent of the planet, with as much as 13.5 Earth masses. Obviously the densities of the atmosphere, mantle, and core are different, just as pressures and temperatures in these three distinctly different regions vary.

The mantle is icy in the sense that it is a hot, dense, electrically conducting mixture of water, ammonia, and less abundant volatile substances. In the mantle, under the great pressures present, molecules dissolved in the water become ionized and therefore create the high electrical conductivity displayed by the mantle. Because of its combined fluid and highly conductive nature, this layer is often referred to as a water-ammonia ocean. The latter quality of the mantle is responsible for generating the planet's complex magnetic field.

The core is far denser, at approximately 9 grams per cubic centimeter. The planet's central pressure and temperature are believed to reach 8 million bars and 5,000 kelvins, respectively. High internal heat flows out to the atmosphere. Materials that have high electrical conductivities usually also have high thermal conductivities. In the case of Uranus, however, it is obvious, based on atmospheric quiescence, that heat flow from the planet's core to the atmosphere is less than that of Neptune. Whereas Neptune radiates more energy than it intercepts by solar irradiance, Uranus is barely 6 percent greater in the infrared than the solar energy absorbed by Uranus's atmosphere. All models agree that Uranus's heat flow from the interior is only 0.042 watt per square meter. The heat flow from the much less massive and smaller Earth, by comparison, is 0.075 watt per square meter. Uranus's low heat flow and hence cold atmosphere (portions of the troposphere have been recorded at a mere 49 kelvins) could be tied to the planet's bizarre rotational configuration. Most planetary scientists believe that Uranus suffered a catastrophic event early in its history, presumably a collision with a large planet-sized body, in order to be left rotating on its side. During such an impact Uranus lost a great deal of its primordial interior heat. Such heat is left over from the gravitational collapse that created the planet. Not all scientists subscribe to that explanation. Some propose that instead the planet's mantle could be layered by composition in such a way that heat flow toward the atmosphere could be diminished by convective action.

METHODS OF STUDY

The Voyager 2 spacecraft has been the only spacecraft to encounter the Uranus system. Originally intended only to fly by Jupiter and Saturn, the Voyager 2 spacecraft was specially targeted to pass through the Uranian system in early 1986. While Voyager 2 cruised to Uranus, a number of astronomers such as Heidi Hammel trained some of the best earthbound telescopes suitable for planetary studies at Uranus to get a feel for what Voyager would encounter. Those studies provided additional information about Uranus's atmosphere but also hinted that Uranus had some internal heat sufficient to drive cloud dynamics. Earthbound telescopic studies did little, however, to enhance understanding of Uranus's interior. A particularly important discovery during the Uranus encounter that aided in describing Uranus's interior was detection of a planetary magnetic field. The nature of Uranus's magnetic field helped planetary scientists devise models for the planet's interior, models quite different from those for the interiors of both Jupiter and Saturn. Then, in 1989, Voyager 2 fulfilled a similar task at Neptune. Planetary scientists realized that the relative similarities of Uranus and Neptune extended to their interiors.

It took the Voyager 2 spacecraft nearly five years to cross the gulf from the Saturn system to the Uranus system. The encounter phase began on November 4, 1985. A great many fundamental questions were about to be answered. At first only ever-increasingly revealing images were produced by Voyager 2. Scientists eagerly awaited detection of a magnetosphere which would indicate Uranus possessed a magnetic field. That would also tell much about the nature of the planet's interior.

When Voyager 2 indeed picked up radio signals indicating the spacecraft crossed the magnetosphere, it clearly revealed Uranus has a magnetic field. After a jam-packed investigation, Voyager 2 ended the Uranus encounter on January 25, 1986, snapping a farewell image of the crescent planet.

CONTEXT

The planets Mercury through Saturn were known to the ancients. No individual can be credited with their discovery. Uranus, however, is the first planet for which definite records exist indicating when the planet was observed and charted. Herschel was not the first to pay special attention to Uranus. Others had misidentified it as an unnamed star or an unknown comet. Herschel recognized Uranus as a newly discovered planet orbiting the Sun. From that time forward to the dawn of the space age, even as Earth-based telescopes grew in size and resolution, little could be learned substantively about Uranus other than its mass, size, mean distance from the Sun, rotation rate, and that it has a very unusual axial tilt relative to the ecliptic plane. It was realized that its atmosphere was quite different from that of both Jupiter and Saturn, suggesting the Uranus's interior differs from those of its two larger gas giant cousins due to its smaller mass. The interior was suspected to be more akin to that of Neptune, which was the next planet in the outer solar system to be discovered after Uranus.

Indeed, models of the two ice giants, generated based on Voyager 2 results and observations, are quite similar. Both planets are believed to have an outer envelope of molecular hydrogen, helium, and methane. Underneath that both planets have mantles that contain water, methane, and ammonia under conditions of high pressures and temperatures. Beneath that is an icy and rocky core. However, the difference between Uranus's and Neptune's interiors is that Uranus's is less active: The planet does not have as great a heat flow from the interior to drive atmospheric dynamics.

David G. Fisher

FURTHER READING

Burgess, Eric. *Uranus and Neptune: The Distant Giants*. New York: Columbia University Press, 1988. Covers the Voyager 2 spacecraft's mission, technical difficulties, and its encounters with Jupiter, Saturn, and Uranus. Focuses on data collected by Voyager 2 about Uranus. Includes several illustrations and tables. Well written, suitable for the general audience.

Elkins-Tanton, Linda T. *Uranus, Neptune, Pluto, and the Outer Solar System*. New York: Chelsea House, 2006. This book explores the Sun's relationship with the three outer planets and their moons. It looks at these planets as recorders of the formation of the solar system. Aimed at a general or high school audience. Illustrations, bibliography, index.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that had investigated all planets by the publication date. Takes an astrophysical approach to place our solar system in a wider context as just one member of similar systems throughout the universe.

Hunt, Garry E., and Patrick Moore. *Atlas of Uranus*. New York: Cambridge University Press, 1988. This was the first volume after the 1986 Voyager encounter to offer a comprehensive history of Uranus: its discovery, satellites, rings, and the data returned by Voyager, including photographs.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs.

Loewen, Nancy. *The Sideways Planet: Uranus*. Mankato, Minn.: Picture Window Books, 2008. An educational children's book devoted to the planet Uranus. Covers Uranus's rings, moons, and tilted axis.

Miner, Ellis. *Uranus: The Planet, Rings, and Satellites*. New York: Ellis Horwood, 1990. The author thoroughly covers the topics of both the Uranian system and the Voyager mission. Illustrations, bibliography, index.

Schmude, Richard W. *Uranus, Neptune, and Pluto and How to Observe Them*. New York: Springer, 2008. Ideal for backyard or amateur astronomers who are interested in observing the outer planets. Also includes up-to-date information about the planets.

Tocci, Salvadore. *A Look at Uranus*. New York: Franklin Watts, 2003. As part of the Out of This World series, this book covers all aspects of the planet Uranus from its discovery through 2002. Includes several photographs. Suitable for all readers.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Eclipses; Infrared Astronomy; Jovian Planets; Planetary Atmospheres;

Planetary Formation; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Telescopes: Ground-Based; Uranus's Atmosphere; Uranus's Magnetic Field; Uranus's Rings; Uranus's Satellites; Uranus's Tilt.

Uranus's Magnetic Field

Categories: Planets and Planetology; The Uranian System

Uranus is the seventh planet outward from the Sun. Although a Jovian or gas giant planet, it has more in common with Neptune than Jupiter or Saturn. Uranus's magnetic field is believed to be produced in a manner similar to that which generates Neptune's magnetic field.

OVERVIEW

Uranus was the first planet to be discovered through telescopic observations. In March, 1781, while searching for binary stars using a 2-meter telescope, William Herschel noted an object he initially thought was either a comet or a nebula. Recognizing its observed orbital motion to be that of a planet, Herschel first proposed the object be named in honor of King George III of England. However, the planet was finally named after the Roman god of the heavens.

Observations of Uranus continued over the next two centuries, but by and large Uranus remained an enticing mystery. Five satellites were discovered between 1787 and 1948. The rotational axis of the planet proved extremely surprising. Uranus is a world rotating virtually on its side. The axis is inclined 97.8° relative to the ecliptic plane of the solar system. Since Uranus's atmosphere—as seen from Earth-based telescopes between the time of Herschel and the dawn of the space age—did not reveal significant features to observe over time, it was not until the second half of the twentieth century that this planet's rotational period was accurately determined.

Knowledge of the orientation of Uranus's rotational axis, its interior structure, and its rotational rate is key to determining the nature of Uranus's magnetic field. Indeed, until the Voyager 2 spacecraft encountered the planet close up in January, 1986, there was no definitive proof that Uranus even possessed a magnetic field, although one was strongly suspected based on a contemporary model of the planet that shared similarities with gas giants Jupiter and Saturn.

Voyager 2 carried a magnetometer and a radio astronomy experiment used in concert to sample the magnetic environment of each planet (Jupiter, Saturn, Uranus, Neptune) that it encountered on its historic "Grand Tour." Basically, a sophisticated magnetometer operates on the very fundamental principal of an induced voltage being produced in a coil of wire when it intercepts a time-varying magnetic flux. That flux is directly proportional to the instantaneous strength of the magnetic field. Spacecraft such as Voyager 2 detect and investigate magnetic environments in space by picking up radiation in radio wavelengths produced by oscillating charged particles, such as those in the solar wind or ions trapped in planetary magnetic fields.

Uranus and Neptune were considered to be gas giants like Jupiter and Saturn until it was realized, based largely upon computer models and Voyager 2 data, that being smaller in mass and having significant atmospheric differences from Jupiter and Saturn, Uranus and Neptune were better classified as ice giants. Finding out about Uranus's interior would be key to determining the means whereby a magnetic field could be produced by the planet. However, the converse is true as well. Directly measuring the magnetic field of the planet would help to develop a model of Uranus's interior. What was known about Uranus's composition in the atmosphere prior to the Voyager 2 encounter was that the planet is composed primarily of hydrogen (83 percent) and helium (15 percent). However, Uranus has a pale blue-green color due to the presence of methane (2 percent), which selectively absorbs red wavelengths of light. Uranus also contains ices such as water, ammonia, and an assortment of hydrocarbons.

KNOWLEDGE GAINED

Voyager 2 determined that the magnetic field generated by Uranus has quite unusual characteristics. The planet's dipole moment is approximately fifty times that of Earth. The value of Uranus's dipole moment is 3.8×10^{17} telsa meters cubed. By comparison to giant Jupiter, this is a mere 0.26 percent of the dipole moment of the largest planet in the solar system. Uranus's average magnetic field strength, the maximum value or amplitude of the field, at the plant's "surface" was measured to be 23 micro-teslas. However, magnetic field strength varied with latitude. At the "surface" in the southern hemisphere, the field strength was seen to dip as low as 10 micro-teslas. Then again, at the "surface" in the northern hemisphere the field was found to be as strong at 110 micro-tesla. This situation contrasts greatly with Earth, where the magnetic field is nearly as intense at both poles. Earth's field is centered close to the planet's physical center. However, that is not the case with Uranus. The center of Uranus's magnetic field actually is displaced from the physical center of the planet by about a third of its radius; the magnetic center is closer to the south rotational pole. Further complicating the field is the fact that the magnetic axis, the line from the south to the north pole through the planet, is tilted 59° relative to the line running between the north and south rotational poles.

Uranus's magnetosphere thus displays a highly unusual tilt. That tilt and the rotational tilt of the planet give the dynamic behavior of Uranus's magnetosphere a twisting structure. To make the magnetosphere's character even stranger, it appears that the ring system around Uranus actually streams ions in the magnetosphere down into Uranus's atmosphere. Auroras are produced, but they differ somewhat from the familiar auroral displays seen in Earth's polar regions.

By detecting variations in radio waves produced by the planet's magnetic field, astronomers have detected the rotational rate of the planet more accurately than by following the differential rotation of those atmospheric features that could be found. That rotation value is 17.233 hours.

Despite the unusual orientation and subsequent twisting behavior of the field lines as the planet rotates, Uranus's magnetosphere does share some things with other planetary magnetic fields in our solar system. Its magnetosphere is affected by the solar wind, forming a bow shock ahead of the planet, a magnetopause, and a magnetotail. The bow shock was crossed at a distance equivalent to 23 Uranus radii, whereas the magnetopause was determined by Voyager 2 to be located at 18 Uranus radii. The magnetotail appears as a corkscrew structure due to the twisting of the planet's magnetic field lines. This magnetic field structure also has given Uranus radiation belts of trapped charged particles. Those particles consist primarily of protons and electrons but have a minor component of molecular hydrogen ions. Uranian satellites create gaps in the radiation belts by "sweeping up" charged particles as they revolve about that planet. Ring particles and the surfaces of the planet's satellites struck by this ionizing radiation are darkened by that exposure.

What generates such an unusual planetary magnetic field? Whereas Earth's field is created deep in the planet by a dynamo effect involving electrical currents generated by its molten core, Uranus's magnetic field is speculated to be produced by the ice giant's mantle. Between the atmosphere and the planet's core, Uranus has a mantle layer believed to be composed of a highly pressurized water, ammonia, and other ices that become ionized under that tremendous pressure and in the presence of temperatures in excess of 1,000 kelvins. Therefore, currents flow through the mantle. Some scientists do not accept this explanation, but at present there is no way, short of direct investigation of the planet's interior, to validate the mantle "ocean" hypothesis or show it to be incorrect.

Why does Uranus's magnetic field have such an unusual orientation relative to the planet's rotational axis? Here there are two reasonable hypotheses. One suggests that the planet is in the process of reversing its magnetic field. (Rocks on Earth present a record that Earth's magnetic field has reversed itself many times over geologic time.) The second hypothesis is that the disruption of the magnetic field alignment resulted from a collision between Uranus and one or more large bodies. Further study will be needed to decide between these two theories or replace them with a better explanation.

The Hubble Space Telescope in concert with the Keck telescope has imaged new storms in Uranus's atmosphere. If Hubble received another servicing mission to extend its life sufficiently, Hubble and Keck would, in addition to searching for atmospheric dynamics, produce data that could assist in developing a better understanding of Uranus's interior. Such data could help explain Uranus's complex magnetic field without getting direct measurements of the field characteristics, as would be possible only by sending another spacecraft to sample the field close to the planet.

CONTEXT

Every 176 years, planetary alignments are such that it is possible through ingenuous use of gravity-assist (or "slingshot") maneuvers to send a pair of spacecraft to visit all the outer planets from Jupiter to Pluto. Prior to the authorization for Mariner Jupiter-Saturn, which later was named Voyager, the National Aeronautics and Space Administration (NASA) had originally proposed sending a pair of sophisticated spacecraft on a journey that had been called the Grand Tour, to take advantage of this rare opportunity to reach four planets. However, that ambitious plan was not funded. The end result was that Voyager 2 would be the only spacecraft to visit Uranus and Neptune in the twentieth century, and most likely for many decades to come after that initial spacecraft encounter. As such, Voyager 2 has provided the greatest share of data about the Uranian system.

The interpretation of those data led to our current understanding about this unique and in many ways bizarre planet. Despite the tremendous insights provided by the Voyager 2 data, many questions remain unanswered. Among those are important questions concerning the internal structure of the planet, the production of the planet's magnetic field, and the nature of the relatively minor amount of internal heating in the planet.

Just as the Galileo probe orbited Jupiter for a prolonged period of time and the Cassini spacecraft did the same at Saturn, the next logical step for Uranus studies would be to dispatch a dedicated orbiter, a spacecraft outfitted with a wide-ranging suite of scientific instruments to allow focused investigations of the planet's atmosphere, internal structure, magnetic field, ring system, and collection of satellites. However, nearly three decades after the Voyager 2 encounter no such program was on the horizon and no proposal was considered likely to be funded unless a nuclear propulsion capability was developed to lessen the travel time to a planet as far away from Earth as Uranus. In the meantime, Uranus will continue to be studied using the Hubble Space Telescope and ground-based facilities (such as the Keck telescope) in attempts to gain further insight into the many unanswered questions remaining from the Voyager encounter.

David G. Fisher

FURTHER READING

Bredeson, Carmen. *NASA Planetary Spacecraft: Galileo, Magellan, Pathfinder, and Voyager*. New York: Enslow, 2000. A part of Enslow's Countdown to Space series, this volume provides an overview of NASA planetary exploration during the last two decades of the twentieth century. Designed for younger readers, but suitable for all audiences.

Burgess, Eric. *Uranus and Neptune: The Distant Giants*. New York: Columbia University Press, 1988. Covers the Voyager 2 spacecraft's mission, technical difficulties, and encounters with Jupiter, Saturn, Uranus, and Neptune. Describes data collected by Voyager 2 about Uranus. Includes several illustrations and tables. Well written and suitable for a general audience.

Elkins-Tanton, Linda T. *Uranus, Neptune, Pluto, and the Outer Solar System*. New York: Chelsea House, 2006. This book explores the Sun's relationship with the three outer planets and their moons. It looks at these planets as recorders of the formation of the solar system. Aimed at a general or high school audience. Illustrations, bibliography, index.

Encrenaz, Thérèse, et al. *The Solar System*. New York: Springer, 2004. A thorough exploration of the solar system from early telescopic observations through the space missions that had investigated all planets by the publication date. The astrophysical approach gives our solar system a wider context as just one member of similar systems throughout the universe.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college text on astronomy, somewhat more advanced than many introductory texts, but with a wealth of detail and excellent diagrams. Chapters 6 through 16 describe the solar system. Comes with a CD-ROM.

Hunt, Garry E., and Patrick Moore. *Atlas of Uranus*. New York: Cambridge University Press, 1988. This was the first volume after the 1986 Voyager encounter to offer a comprehensive history of Uranus: its discovery, satellites, rings, and the data returned by Voyager, including photographs.

Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006. Suitable as a textbook for upper-level college courses in planetary science. Focuses on Jupiter, Saturn, Uranus, and Neptune and their satellites, rings, and magnetic fields. Filled with figures and photographs. Accessible to the serious general audience.

Loewen, Nancy. *The Sideways Planet: Uranus*. Mankato, Minn.: Picture Window Books, 2008. An educational children's book devoted to the planet Uranus. Covers the planet's rings, moons, and tilted axis.

McBride, Neil, and Iain Gilmour, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004. A complete description of solar system astronomy suitable for an introductory college course but useful to interested laypersons as well. Filled with supplemental learning aids and solved student exercises. A Web site is available for educator support.

Miner, Ellis. *Uranus: The Planet, Rings, and Satellites*. New York: Ellis Horwood, 1990. The author thoroughly covers the topics of both the Uranian system and the Voyager mission. Illustrations, bibliography, index.

Schmude, Richard W. *Uranus, Neptune, and Pluto and How to Observe Them*. New York: Springer, 2008. Ideal for backyard or amateur astronomers who are interested in observing the outer planets. Includes up-to-date information about the planets.

Tocci, Salvadore. *A Look at Uranus*. New York: Franklin Watts, 2003. As part of the Out of This World series, this book covers all aspects of the planet Uranus from its discovery through 2002. Includes several photographs. Suitable for all readers.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Eclipses; Infrared Astronomy; Jovian Planets; Planetary Atmospheres; Planetary Formation; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Telescopes: Ground-Based; Uranus's Atmosphere; Uranus's Interior; Uranus's Rings; Uranus's Satellites; Uranus's Tilt.

Uranus's Rings

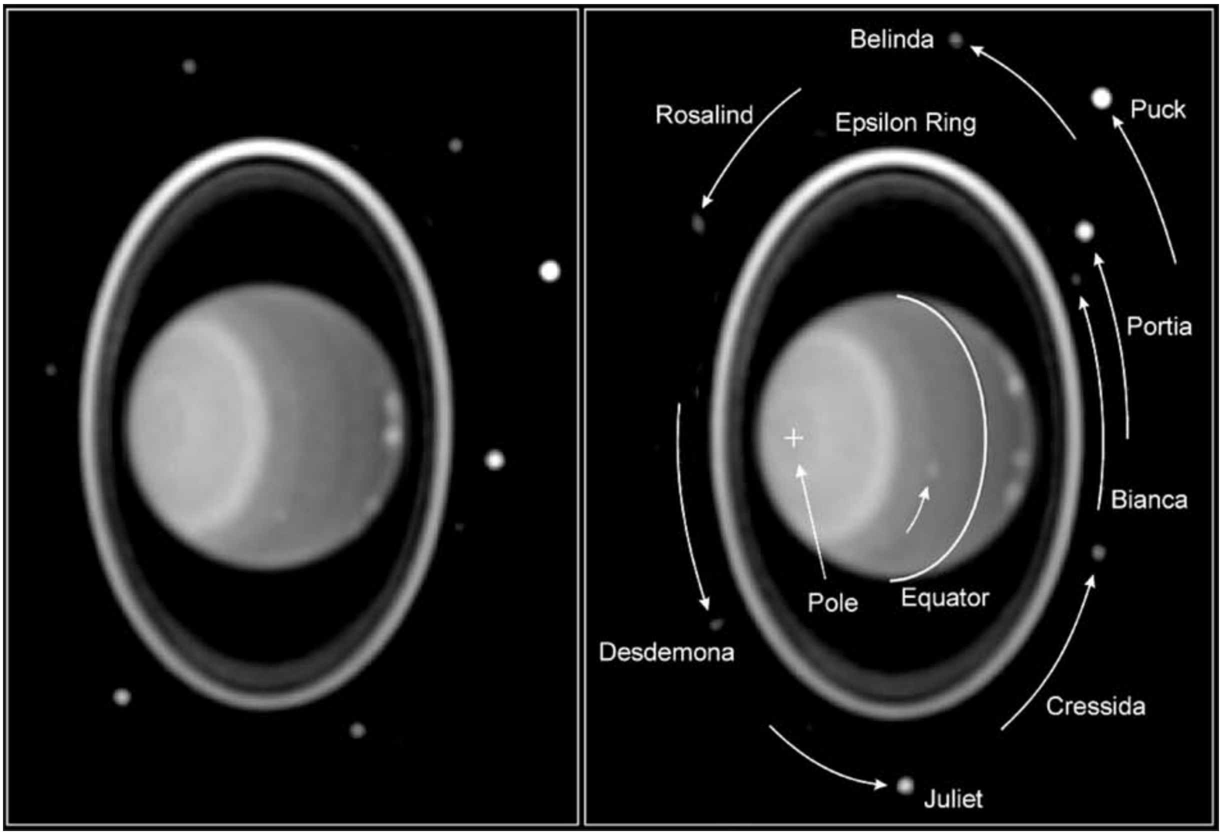
Categories: Planets and Planetology; The Uranian System

Uranus has thirteen known rings, eleven inner and two outer ones. These rings are mostly very narrow and faint. Uranus's ring system is less complex than Saturn's but more so than that of Jupiter.

OVERVIEW

When Uranus was first observed (perhaps as early as 1690), many astronomers considered it to be a star rather than a planet. William Herschel studied Uranus in 1781, when he thought he had discovered a new comet. After two years of further study, astronomers agreed that Uranus was in fact a planet.

Herschel appears to have observed the rings of Uranus in February of 1789. He sketched an image of Uranus in his journal, making a note that the planet had rings of a faint reddish hue. Rings around Uranus remained an open issue for a long time. The existence of Uranus's ring system was finally confirmed, albeit accidentally, in 1977. Astronomers James Elliot, Ed-



This 1997 image from the Near Infrared Camera and Multi-Object Spectrometer (NICMOS) on the Hubble Space Telescope shows clouds on Uranus (left) and clearly displays the planet's rings, satellites, and rotation (right). (NASA)

ward Dunham, and Douglas Mink set out to study Uranus's atmosphere by observing the occultation of the star SAO 158687. They noticed that this star briefly disappeared from view before and after being eclipsed by the planet. Each of the five occultations that they observed yielded the same results. When the group published their work, they referenced these occultations and resulting rings using the Greek letters Alpha, Beta, Gamma, Delta, and Epsilon. In 1978 another group of scientists found four additional rings. The Eta ring was found between the Beta and Gamma rings. The other three were discovered inside the orbit of the Alpha ring and were named Six, Five, and Four (in that order).

The National Aeronautics and Space Administration's (NASA's) robotic Voyager 2 probe flew by Uranus in 1986. The spacecraft took the first photographs of the Uranian ring system. Voyager 2 also discovered two more faint rings, Lambda and 1986U2R/Zeta, bringing the total number of known rings around the planet to eleven. In 2003, the Hubble Space Telescope discovered, and in 2005 confirmed, the existence of an additional pair of rings. They form an outer ring system that is separate from the other eleven rings. These two outer rings have an orbital radius of more than 100,000 kilometers from Uranus's center—double that of the inner rings. In addition to the twelfth and thirteenth rings, Hubble found two satellites. Mab, which is only 24 kilometers in diameter, shares an orbit with the outermost ring. Every time a meteoroid impacts the small satellite, dust particles and other debris ejected become part of the Mu ring. The Nu ring lies between Uranus's small satellites Rosalind and Portia.

Rings of Uranus

	Radius (km)	Radius/ Eq. Radius	Optical Depth	Albedo ($\times 10^{-3}$)	Width (km)	Eccentricity
Uranus equator	25,559	1.000	—	—	—	—
6	41,837	1.637	~0.3	~15	1.5	0.0010
5	42,234	1.652	~0.5	~15	~2	0.0019
4	42,571	1.666	~0.3	~15	~2	0.0011
Alpha	44,718	1.750	~0.4	~15	4-10	0.0008
Beta	45,661	1.786	~0.3	~15	5-11	0.0004
Eta	47,176	1.834	~0.4-	~15	1.6	—
Gamma	47,627	1.863	~1.3+	~15	1-4	0.0011
Delta	48,300	1.900	~0.5	~15	3-7	0.00004
Lambda	50,024	1.957	~0.1	~15	~2	0.
Epsilon	51,149	2.006	0.5-2.3	~18	20-96	0.0079

Source: Data are from the National Aeronautics and Space Administration/Goddard Space Flight Center, National Space Science Data Center.

The Mu and Nu rings are very different from the inner rings. With widths of 17,000 kilometers and 30,000 kilometers, respectively, the Mu and Nu rings are much broader. These two rings are also much fainter then the others, but they can be seen on the Voyager 2 photographs.

Many similarities exist between Uranus's outer rings and Saturn's E and G rings. The E ring includes Enceladus, which contributes dust to it the same way Mab is believed to contribute to the Mu ring. The Nu ring, like Saturn's G ring, contains no embedded "shepherding" satellites and is composed of dust and larger particles. Scientists working with the Keck telescopes in Hawaii studied the rings at near-infrared wavelengths. The Nu ring was visible, meaning it has a reddish hue. This possibly gives some credit to Herschel's original claim about observing Uranus's rings, despite critics' claims that the rings are too faint for him to have seen. The Mu ring was not visible, meaning that its small dust particles appear blue in color. Red is a typical color for planetary rings. Blue however, is not. Saturn's E ring is the only other ring known to have the unusual blue hue.

The inner ring system contains two types of rings: narrow and dusty. The closest ring to Uranus is 1986U2R. It was discovered in 1986 by Voyager 2. This ring is only about 12,000 kilometers above the cloud tops of Uranus. The 1986U2R (or Zeta) ring was observed in 2003 and 2004 using the Keck telescopes. Scientists found the ring to be broad, very faint, and composed of dust grains.

The next set of rings is Six, Five, and Four, which were named for the occultations that led to their discoveries. They are the faintest of Uranus's narrow main rings. These three lie outside Uranus's equatorial plane by 0.06°, 0.05°, and 0.03°, respectively. Six, Five, and Four do not contain dust and are the thinnest of Uranus's narrow rings.

After the Epsilon ring, the Alpha and Beta rings are the brightest of Uranus's rings. Alpha and Beta are narrowest and faintest at their closest points to Uranus. At their farthest, the two rings are their broadest and brightest. Like all of Uranus's rings, Alpha and Beta are composed of extremely dark material. They are much darker than Uranus's inner satellites, meaning that the rings cannot be composed of pure water ice. The composition of the rings is thus unknown, but astronomers think it is probably a mixture of dark materials and ices.

The seventh ring outward from Uranus's core is Eta, at 47,176 kilometers. Eta has both a narrow part and a broader, dustier section. When Voyager 2 photographed the ring in forward scattered light, it appeared very bright, indicating a large amount of dust. Eta has an inclination and eccentricity of zero, meaning that the ring lies in the planet's equatorial plane, and the ring particles execute circular orbits. In 2007 Uranus's rings were viewed edge-on for the first time. The Eta ring appeared to be the second brightest, which was a significant increase. This finding has led planetary scientists to believe that, while the ring is optically narrow, it is geometrically thick. The next chance for astronomers to view the rings in that unique geometry will occur in 2049.

The Gamma ring is narrow, with an inclination close to zero. Gamma's width varies from 3.6 to 4.7 kilometers. The ring was not visible during the 2007 ring plane crossing. This means that Gamma is both optically and geometrically narrow. The ring also does not contain any dust. Scientists are uncertain about what holds this small ring together.

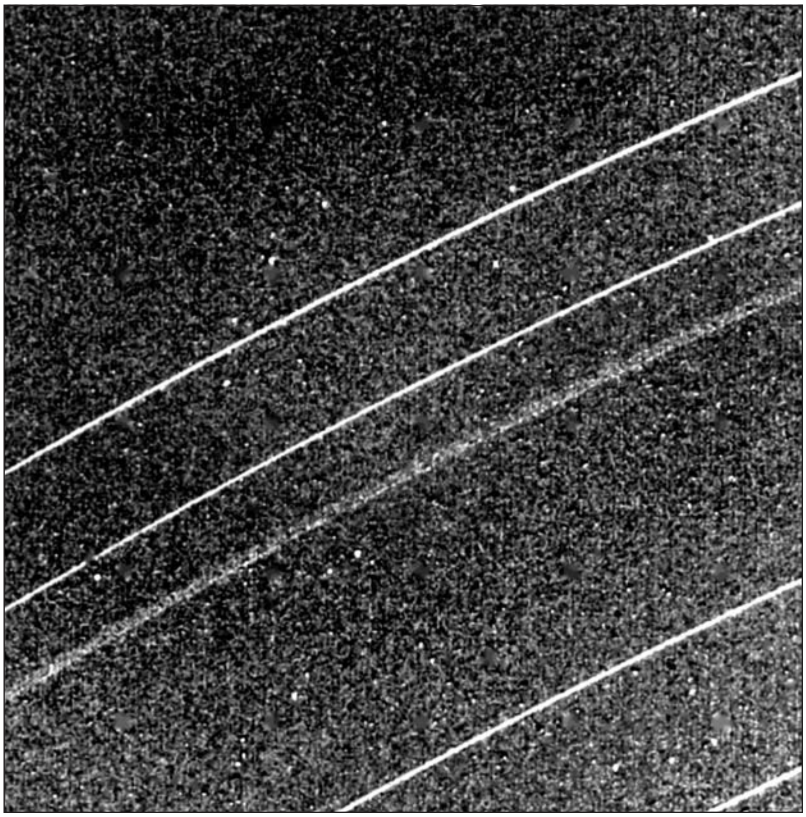
Like the Eta ring, Delta has both a narrow and a broad component. The thinner part varies from 4.1 to 6.1 kilometers wide, while the thicker part ranges from 10 to 12 kilometers. The wide section is composed of dust, unlike the narrow part. In 2007, only the broad area of Delta was visible. At the outer edge of the Delta ring, a small satellite named Cordelia orbits Uranus.

The Lambda ring lies between Cordelia (a shepherd satellite) and the Epsilon ring. Lambda is faint and narrow even when backlit. The dusty ring was first detected by Voyager 2 during stellar occultation observations, but only at ultraviolet wavelengths. The Lambda ring is composed of micrometer-sized dust, which was confirmed in 2007, when it appeared very bright.

The brightest and densest of Uranus's rings is Epsilon. It is the outermost ring of the inner system. Epsilon reflects two-thirds of all the light visible from Uranus's rings. It is only one of two rings that Voyager 2 was able to photograph clearly. Epsilon is the most eccentric of the rings but has a near-zero orbital inclination. The dense ring has particles ranging in size from 0.2 to 20 meters in diameter. In 2007, the ring was not observable, because of its lack of dust. Epsilon contains many dense, narrow ringlets, and possibly partial arcs. The ring may stay so compact because of its shepherd satellites: Cordelia on the inner side and Ophelia on the outer.

KNOWLEDGE GAINED

The first rings of Uranus were officially discovered by accident in 1977. Elliot, Dunham, and Mink were using the Kuiper Airborne Observatory (KAO) to study Uranus's atmosphere during five stellar occultations. The KAO is an



Voyager 2 took these images of Uranus's rings in January, 1986, from a distance of 1.12 million kilometers. From the top, the Delta, Gamma, Eta, Beta, and Alpha rings are visible. (NASA/JPL) airplane with a 36-inch (91.5-centimeter) telescope mounted on the side. With the KAO, scientists can conduct research while flying 14 kilometers above the Earth's surface, thereby making infrared observations readily. At that elevation there is significantly less atmospheric water vapor, which blocks infrared wavelengths from reaching the surface of the Earth. The KAO therefore combines many benefits of a space telescope with the accessibility of a ground-based telescope.

Launched in 1977, Voyager 2 is the only spacecraft to have visited Uranus. It came within 81,500 kilometers of the planet on January 24, 1986. The spacecraft had several instruments on board, including cameras, magnetometers, and spectroscopes. At the time, Uranus's south pole was pointed toward the Sun. Voyager discovered ten satellites as well as the Lambda and Zeta (1986U2R) rings. The Mu and Nu rings have since been located on Voyager 2 photographs.

The Hubble Space Telescope was studying Uranus in 2003 when it discovered the Mu and Nu rings. Scientists were able to confirm the finding in 2005. When the Keck telescope studied the two rings at near-infrared wavelengths, only the Nu ring was visible. This means that the Nu ring has a reddish color. The Mu ring therefore has a bluish tint, because it was not visible.

In 2007, astronomers were able to view Uranus's rings edge-on. Teams of scientists used the Keck II telescope in Hawaii, the Hubble Space Telescope, and the European Southern Observatory's Very Large Telescope in Chile to study the event. Images taken with the Keck telescope show that the rings have changed since Voyager 2 visited the planet more than two decades ago. The broad, dusty inner Zeta ring appears very different. If it is the same ring discovered by Voyager, Zeta has moved several thousand kilometers away from Uranus. Similar shifts have been detected in the ring systems of Saturn and Neptune.

CONTEXT

All of the Jovian planets in our solar system have ring systems. Each set of rings is unique. Neptune's rings are simpler than Uranus's, containing only five rings and partial arcs. Saturn's ring system on the other hand is more complex. Jupiter has but two faint rings.

The thirteen rings of Uranus are mostly faint and narrow. They are in two groupings—an inner system of eleven rings and an outer set of two. In 2007, the Nu ring was determined to be red in color, and the Mu ring was found to be blue. Red seems to be a typical color for planetary rings, like Saturn's G ring. The blue color of the Mu ring, however, is not common. The only other example in the solar system is Saturn's E ring. What caused this odd blue color is still a mystery to scientists. The rings around Uranus are believed to be made of debris from collisions between Uranus's satellites.

Scientists can learn more about the formation and evolution of the solar system by investigating planetary ring systems. As ground-based and space telescopes improve, astronomers could unlock the secrets of Uranus's rings and the solar system itself.

Jennifer L. Campbell

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A well-written college-level text for introductory astronomy courses. Has a chapter on Uranus and Neptune that covers the ring systems.

Elkins-Tanton, Linda T. *Uranus, Neptune, Pluto, and the Outer Solar System*. New York: Chelsea House, 2006. Explores the Sun's relationship with the three outer planets and their moons, considering these planets as recorders of the formation of the solar system. Aimed at a general or high school audience. Illustrations, bibliography, index.

Esposito, Larry. *Planetary Rings*. New York: Cambridge University Press, 2006. A synopsis of current knowledge of the outer planets' ring systems. Includes information from the Cassini mission on Uranus's rings, and on ring ages and evolution. Geared toward scientists and college students.

Fridman, Alexei M., and Nikolai N. Gorkavyi. *Physics of Planetary Rings: Celestial Mechanics of Continuous Media*. New York: Springer, 1999. Compares the ring systems of Jupiter, Saturn, Uranus, and Neptune using observational and mathematical data. Designed for scientists, astronomy and physics students, and amateur astronomers wishing to know more about the rings of the outer planets.

Hunt, Garry E., and Patrick Moore. *Atlas of Uranus*. New York: Cambridge University Press, 1988. This was the first volume after the 1986 Voyager encounter to offer a comprehensive history of Uranus: its discovery, satellites, rings, and the data returned by Voyager, including photographs.

Loewen, Nancy. *The Sideways Planet: Uranus*. Mankato, Minn.: Picture Window Books, 2008. An educational children's book devoted to the planet Uranus. Covers Uranus's rings, moons, and tilted axis.

Miner, Ellis. *Uranus: The Planet, Rings, and Satellites*. New York: Ellis Horwood, 1990. The author thoroughly covers the topics of both the Uranian system and the Voyager mission. Illustrations, bibliography, index.

Miner, Ellis D., Randii R. Wessen, and Jeffrey N. Cuzzi. *Planetary Ring Systems*. New York: Springer Praxis, 2006. Looks at the ring systems of each gas giant. Covers recent research in the field, as well as the many questions that remain unanswered.

Schmude, Richard W. *Uranus, Neptune, and Pluto and How to Observe Them*. New York: Springer, 2008. Ideal for backyard or amateur astronomers who are interested in observing the outer planets. Includes up-to-date information about the planets.

Tocci, Salvadore. *A Look at Uranus*. New York: Franklin Watts, 2003. As part of the Out of This World series, this book covers all aspects of the planet Uranus, including the planet's discovery and research to 2002. Includes several photographs. Suitable for all readers.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Eclipses; Infrared Astronomy; Jovian Planets; Planetary Atmospheres; Planetary Formation; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Telescopes: Ground-Based; Uranus's Atmosphere; Uranus's Interior; Uranus's Magnetic Field; Uranus's Satellites; Uranus's Tilt.

Uranus's Satellites

Categories: Natural Planetary Satellites; Planets and Planetology; The Uranian System

Uranus's natural satellites form a miniature solar system with distinctive properties that teach us the complexity and diversity of planetary and satellite formation. The peculiar surface features of some of the satellites and their unusual orbital characteristics suggest an earlier epic in the solar system with violent collisions among its members.

OVERVIEW

Uranus holds its place in the solar system as a member of the subgroup of planets that are called Jovian, after the largest planet in the group, Jupiter. These planets are also referred to as "gas giants" in that their atmospheres comprise the greatest portion of the planet's structure. It is the physical nature and orbital properties of the satellites that set apart one Jovian planet from another, and Uranus is no exception. Surveying the most interesting properties of this planetary system and highlighting their respective features can provide insights about the origin and evolution of the solar system itself.

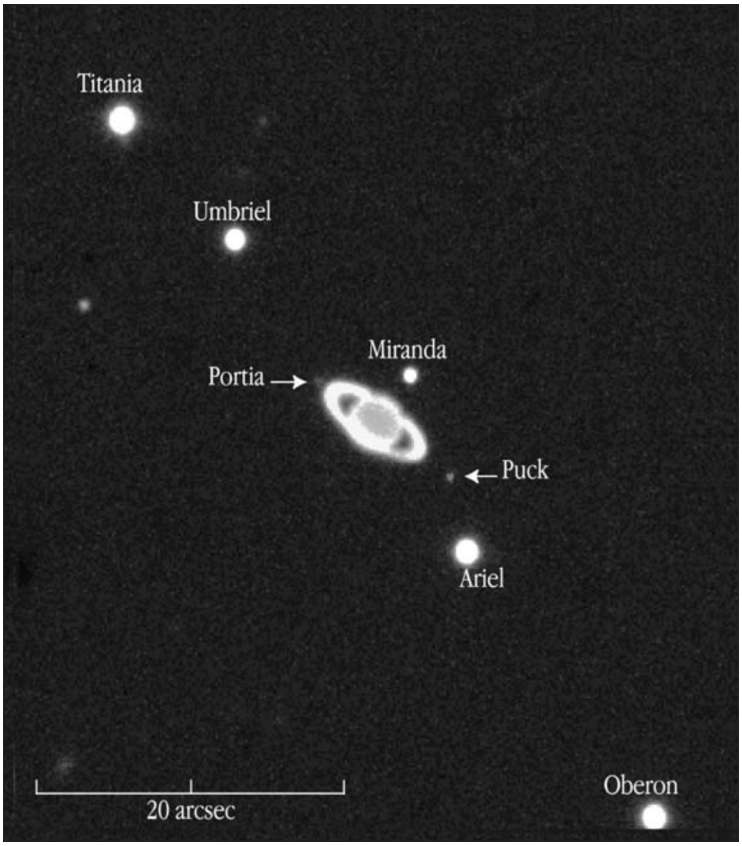
Uranus has twenty-seven satellites that have been identified. Their names follow a theme that is distinctive in the solar system in that the satellites are not named for mythological figures, like those of the other planets, but instead take their names from characters in plays by William Shakespeare and poems by Alexander Pope. Oberon, Titania, and Puck, for example, were named for characters from Shakespeare's *A Midsummer Night's Dream*; Ariel, Umbriel, and Belinda are named for characters in Pope's *The Rape of the Lock*. Oberon and Titania were first discovered by William Herschel in 1781. Ariel and Umbriel were discovered in 1851 by William Lassell. In 1948, Gerard Kuiper discovered the last moon of any significant size, Miranda.

The only spacecraft to visit Uranus to date has been Voyager 2, which flew by the planet in

1986. Despite the briefness of its visit, Voyager 2 discovered ten small satellites: Juliet, Puck, Cordelia, Ophelia, Bianca, Desdemona, Portia, Rosalind, Cressida, and Belinda. (Perdita was also imaged by Voyager 2, but its discovery was not confirmed until 1999.) The number of satellites has swelled to the current known number of twenty-seven through observation with both orbiting telescopes (Hubble Space Telescope) and ground-based observatories.

A way to organize the Uranian systems of satellites is to think of them as distributed in three divisions. The first division consists of the inner thiteen, relatively small, circular satellites starting just outside the ring system. The second division comprises the five midsized satellites that were discovered prior to the Voyager 2 mission. Finally, the last division comprises nine irregular satellites discovered more recently.

None of the natural satellites of Uranus can be considered on the scale of the largest satellites of the other planets of the solar system, such as Titan (Saturn), Triton (Neptune), the four Galilean satellites of Jupiter (Io, Europa, Callisto, and Ganymede), or even Earth's own Moon. Nor does any of these satellites contain an atmosphere, as does as Titan, or show active volcanoes, as seen on Io. However, the five largest Uranian satellites in Division 2—Miranda, Ariel, Umbriel, Titania, and Oberon (in order outward from the planet)—all have enough mass to be spherical. Hence, if they were not orbiting Uranus and were free from any debris, they would qualify as dwarf planets, like Pluto. The composition of these satellites is mostly ice, with mixtures of ammonia and methane. They all exhibit synchronous rotation, rotating exactly at the same rate they revolve around Uranus, always showing the same side facing the planet.



A near-infrared image of the moons of Uranus from the European Southern Observatory, 2002. (European Southern Observatory)

The large satellites of Uranus all show extensive cratering. The consensus of scientific opinion is that the larger craters were formed from collisions with planet-sized objects during the formation of the solar system, and the smaller craters were produced afterward from impacts from comets and meteoroids. Ariel, Umbriel, and Miranda have unusual surface features. Ariel, the lightest-colored, has craters, valleys, and canyons. It has a diameter of about 1,300 kilometers with grooves and crevices that extend over its entire surface, suggesting recent lava flow that has cooled and solidified. Umbriel is the darkest-colored satellite; its surface is old and cratered. Although Umbriel (with a diameter of 1,110 kilometers) shows the most uniform cratering, with little evidence of geological activity, it has a large, bright ring at the top edge of its southern hemisphere; this ring's origin is unknown, but it has unofficially been named the "fluorescent Cheerio." Miranda (about 500 kilometers diameter), Major Uranian Satellites

Satellite	*Diameter (km)*	*Distance from Planet (km)*
Miranda	500 ± 220	130,000
Ariel	1,300 ± 130	192,000
Umbriel	1,110 ± 100	267,000
Titania	1,600 ± 120	438,000
Oberon	1,630 ± 140	596,000

Source: Data are from Jet Propulsion Laboratory, California Institute of Technology. *Voyager at Uranus: 1986*. JPL 400-268. Pasadena, Calif.: Author, 1985, p. 5.

the smallest of these five large satellites, has the most distinctive surface features. There are sharp grooves and ridges along its surface. One feature resembles a large chevron, and another looks like a carved racetrack. Titania and Oberon are both larger than the other satellites (with diameters of approximately 1,500-1,600 kilometers), and they are also several times farther away from Uranus. Both Titania and Oberon show large craters, but the presence of circular regions on Oberon suggests that it has experienced more geological activity than Titania. Oberon's surface is frozen and has a mountain 6 kilometers high. Titania, the biggest satellite among those of Uranus, has impact basins, craters, and rifts.

Inside Miranda's orbit there are the thirteen Division 1 satellites, all with enough mass to be spherical and with diameters ranging from about 25 to 170 kilometers. They are (in order out from the planet) Cordelia, Ophelia, Bianca, Cressida, Desdemona, Juliet, Portia, Rosalind, Cupid, Belinda, Puck, and Mab. Their composition appears to be about half water and half rocklike materials. Their orbital distribution presents a more crowded arrangement than that of the large satellites, and it appears that these thirteen bodies interact gravitationally with each other, crossing paths and periodically colliding with one another. Some of these inner satellites may serve as shepherds for Uranus's narrow rings, which orbit closest to the planet.

Outside the orbit of Oberon are the small irregular satellites in Division 3: Francisco, Caliban, Stephano, Trinculo, Sycroax, Margaret, Prospero, Setebos, and Ferdinand. They are considered irregular, though, because they orbit in odd directions far from Uranus. Their respective sizes and compositions have not been measured accurately. Scientists have estimated that their diameters range from 18 to 150 kilometers. They have very eccentric (elliptical) orbits, and all orbit retrograde, that is, in the direction opposite to the general revolution of other bodies in the solar system, which is counterclockwise around the north polar axis of the Sun.

KNOWLEDGE GAINED

The study of Uranus's system of natural satellites, including their orbital and physical properties, makes it possible to understand more fully the processes not only of planet formation but of the evolution of the solar system itself. Satellites can be formed at the time of the formation of the planet they orbit, but they can also be captured by the planet's gravitational pull at a later stage. It appears that both processes are at work in the Uranian system. The group of satellites that are in Division 2 (Miranda, Ariel, Umbriel, Titania and Oberon), as well as the thirteen satellites in Division 1, most likely formed at the same time Uranus condensed. However, events have transpired to alter these satellites' orbits and surface features. For example, Miranda has a scarred surface that is presumed to be the result of a collision with another object and was fractured into pieces. These pieces then came back together unevenly and show the rough terrain and scarps that characterize the satellite. An alternative hypothesis is that Miranda, being too small an object to complete its internal mixing, froze midway through the process of separating its structure into layers.

Several of the Division 1 moons are in such close orbits with each other that they conceivably collide and switch orbital positions around Uranus. In addition, Cordelia and Ophelia serve as shepherding moons for the Epsilon ring system.

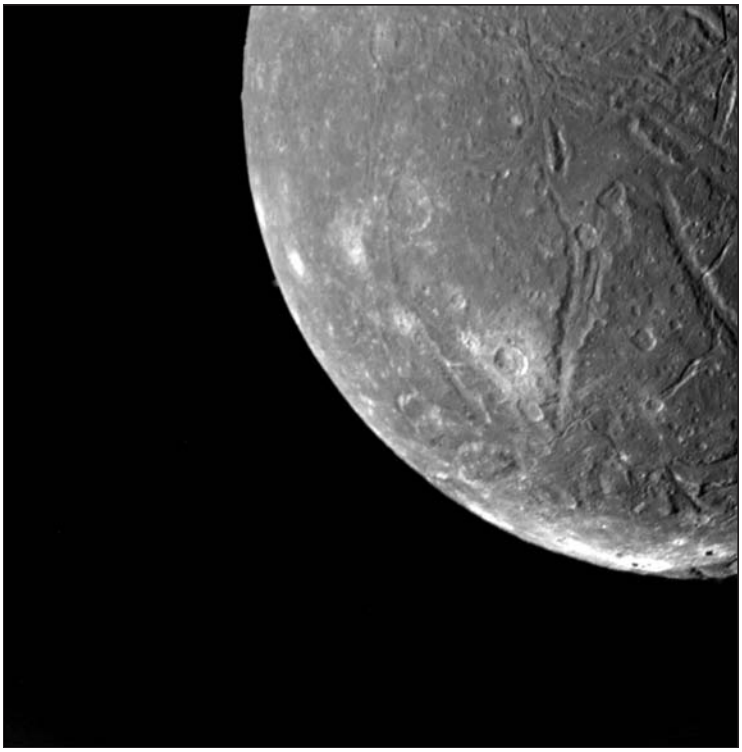
The irregular satellites in Division 3 appear to be satellites that were previously solar system bodies, such as comets and asteroids, that were captured by Uranus's gravitational pull and orbit around the planet in elliptical and retrograde orbits.

This accumulated knowledge points to a solar system that evolved sequentially over vast periods of time with many diverse objects that are still changing their orbital shapes and constitutions.

CONTEXT

There were only five known satellites of Uranus until Voyager 2 visited this planetary system and made its closet approach on January 24, 1986. Voyager 2 was launched in 1977 and visited Jupiter in 1979, Saturn in 1981, and, after its flyby of Uranus, Neptune in 1989.

As Voyager 2 approached the Uranus system, its onboard computers were reprogrammed by scientists and engineers back on Earth at the Jet Propulsion Lab to enable the cameras to produce high-quality photographs in the reduced light and at the high speeds at which the spacecraft would be traveling on its flyby. Most of the photographs were taken in a six-hour period in and around the time of closest approach (9:59 A.M. PST) on January 24. On the way to the rendezvous with Uranus, Voyager 2 obtained clear, high-resolution images of each of the five large Uranian satellites (Miranda, Ariel, Umbriel, Titania, and Oberon). It was its discovery of eleven new satellites that added to our knowledge base of this planet. During processing of images of the outer ring (Epsilon ring) of Uranus, it was discovered that two small satellites, Cordelia and Ophelia, were shepherding or keeping this thin ring in orbit around the planet. Also sighted were Bianca, Cressida, Desdemona, Juliet, Portia, Rosalind, Cupid, Belinda, and Perdita which belong to what has been called the Portia Group of satellites, These satellites have similar orbits and light-reflecting properties. The closeness of their respective orbits leads to the hypothesis that this group interacts with each other and may at times collide.



Light-colored Ariel, seen in this 1986 image from Voyager 2, has grooves and crevices that extend over its entire surface, suggesting recent lava flow that has cooled and solidified. (NASA/JPL)

Until 1997, the Uranian system was distinct from the other Jovian planets in that there were no identified irregular satellites. However, the discovery on September 6, 1997, of two irregular satellites, Caliban and Sycroax, by Brett J. Gladman, Philip D. Nicholson, Joseph A. Burns, J. J. Kavelaars, Brian G. Marsden, Gareth V. Williams, and Warren B. Offutt, using the 200-inch Hale telescope (at Palomar Observatory in Southern California) removed that distinction. Subsequently, Stephano, Prospero, and Setebos were discovered by Matthew J. Holman, Kavelaars, Gladman, Jean-Marc Petit, and Hans Scholl on July 18, 1999. Trinculo, Margaret, and Ferdinand were discovered by Holman, Kavelaars, and Dan Milisavljevic on August 13, 2001.

Joseph Di Rienzi

FURTHER READING

Bennett, Jeffery, Megan Donahue, Nicholas Schneider, and Mark Voit. *The Cosmic Perspective*. 3d ed. San Francisco: Pearson Addison Wesley, 2004. This textbook provides a thematically organized overview of the universe. Chapter 12 discusses the Jovian systems and contains a subsection on the medium-sized satellites of Uranus.

Burgess, Eric. *Uranus and Neptune: The Distant Giants*. New York: Columbia University Press, 1988. Covers the Voyager 2 spacecraft's mission, technical difficulties, and its encounters with Jupiter, Saturn, and Uranus. Describes the data collected by Voyager 2 about Uranus. Includes several illustrations and tables. Well written and suitable for the general audience.

Carroll, Bradley W., and Dale A. Ostlie. *An Introduction to Modern Astrophysics*. San Francisco: Pearson Addison Wesley, 2007. This is an encyclopedic textbook that covers all of modern astronomy and astrophysics. Although much of the book is for the advanced student, the chapters on the solar system are very descriptive. Chapter 21, "The Realms of the Giant Planets," includes a section on their satellites and discusses Miranda in particular.

Elkins-Tanton, Linda T. *Uranus, Neptune, Pluto, and the Outer Solar System*. New York: Chelsea House, 2006. This book explores the Sun's relationship with the three outer planets and their moons. It looks at these planets as recorders of the formation of the solar system. Aimed at a general or high school audience. Illustrations, bibliography, index.

Encrenaz, Thérèse, et al., eds. *The Outer Planets and Their Moons: Comparative Studies of the Outer Planets Prior to the Exploration of the Saturn System by Cassini-Huygens*. New York: Springer, 2005. An in-depth look at the current understanding of the solar system's outer planets. Focuses on the studies of their formation, evolution, magnetospheres, satellites, and ring structures. For scientists, first-year graduate students, and advanced undergraduates.

Hunt, Garry E., and Patrick Moore. *Atlas of Uranus*. New York: Cambridge University Press, 1988. This was the first volume after the 1986 Voyager encounter to offer a comprehensive history of Uranus: its discovery, satellites, rings, and the data returned by Voyager, including photographs.

Loewen, Nancy. *The Sideways Planet: Uranus*. Mankato, Minn.: Picture Window Books, 2008. An educational children's book devoted to the planet Uranus. Covers Uranus's rings, moons, and tilted axis.

Miner, Ellis. *Uranus: The Planet, Rings, and Satellites*. New York: Ellis Horwood, 1990. The author thoroughly covers the topics of both the Uranian system and the Voyager mission. Miranda is featured as a remarkable satellite. Illustrations, bibliography, index.

Schmude, Richard W. *Uranus, Neptune, and Pluto and How to Observe Them*. New York: Springer, 2008. Ideal for backyard or amateur astronomers who are interested in observing the outer planets. Also includes up-to-date information about the planets.

Tocci, Salvadore. *A Look at Uranus*. New York: Franklin Watts, 2003. As part of the Out of This World series, this book covers all aspects of the planet Uranus, from its discovery through 2002. Includes several photographs. Suitable for all readers.

See also: Jupiter's Satellites; Miranda; Neptune's Satellites; Planetary Ring Systems; Planetary Satellites; Saturn's Satellites; Telescopes: Ground-Based; Titan; Triton; Uranus's Magnetic Field; Uranus's Rings.

Uranus's Tilt

Categories: Planets and Planetology; The Uranian System

All planets in our solar system rotate on an axis tilted in relation to the ecliptic plane (the plane carved out by their orbit around the Sun). However, Uranus's axis is tilted at such an extreme angle that the planet rotates while virtually lying on its side. Despite several theories, scientists do not fully understand what causes Uranus's tilt.

OVERVIEW

Uranus was observed as early as 1690, but astronomers thought it was a star. Using a telescope he built himself, Sir William Herschel observed Uranus over a series of nights in 1781. Herschel initially reported to the Royal Society that he had discovered a new comet. After tracking the "comet" for two years, astronomers finally agreed that Uranus was actually the seventh planet in the solar system.

In 1829, astronomers determined that the rotation of Uranus was unique. All of the planets rotate on an axis that is tilted with respect to the orbital plane of the solar system. The orbital plane is the imaginary surface on which the planets orbit and almost lies on the Sun's equator (the plane is tilted at a 7° angle with respect to the Sun's equator). Axial tilt is calculated by drawing a line perpendicular to the orbital plane. The rotational axis of the planet is compared to the perpendicular line. For example, the rotational axis of the Earth has a tilt of 23.5°. Mars is tilted at 25.19°, and Saturn's axis is tilted at 26.73°. Uranus, on the other hand, has an axial tilt of 97.8°. Because of this, Uranus is often referred to as the "sideways planet." Either the north or the south pole of Uranus is usually pointed toward the Sun. Uranus's equator experiences day and night the same way as the Earth's polar ice caps do. The poles of Uranus each experience forty-two years of sunlight, followed by forty-two years of complete darkness. Only around equinoxes is the Sun facing Uranus's equator, causing "normal" Earth day-night conditions. Its last equinox occurred on December 7, 2007, and the the next will not happen until the year 2049.

Due to its unusual orientation, scientists have conflicting methods for determining which pole is "north" and which is "south." The International Astronomical Union (IAU) refers to whichever pole lies above the orbital plane as the north pole. Most scientists use this designation. Others use the right-hand rule from physics and the direction the planet is spinning to designate the poles north or south. This method contradicts the IAU's determination, instead naming the pole below the orbital plane as "north."

The only spacecraft to date that has visited Uranus is Voyager 2. Launched in 1977, the probe reached Uranus in 1986. Voyager 2 came within 81,500 kilometers of the planet. It discovered and photographed ten new satellites and nine rings orbiting Uranus. The spacecraft also helped scientists determine more precisely the axial tilt of Uranus.

There are two main competing theories to explain why Uranus is tilted on its side. No one knows who proposed the popular "collision" theory, which posits that Uranus formed and then a large Earth-sized object crashed into it with such force that it left the planet on its side. The current accepted theory of planetary formation is the idea of nebular condensation, developed in the seventeenth century by French philosopher René Descartes. As a massive cloud of interstellar dust and debris condensed, it would collapse and start to spin. Planets slowly would begin to form from clumps of matter joining together. The bigger the planets grew, the faster they would be able to attract more material through a process known as accretion. The debris cloud that the planets formed from is called the accretion disk, which became the orbital plane. The collision knocking Uranus on its side would have had to happen early in its formation. Possibly an object struck the planet's core before Uranus's satellites had condensed from the debris cloud surrounding it. Another theory is that the impact left behind debris that later became Uranus's satellites. However, there are several questions that remain unanswered with this scenario.

Why does Uranus have a nearly circular obit, like the other planets? Would not a large impact have affected Uranus's orbit? If Uranus's satellites had formed before the collision, why were their orbits not changed? The satellites orbit Uranus's equator, just like its ring system. Two very small captured satellites, however, have been found orbiting Uranus's poles. The nebular theory also fails to explain other oddities of the solar system, such as why Venus has a retrograde rotation (rotates backward), why Mercury and Pluto have elliptical orbits, and why Uranus and Neptune have tilted magnetic fields.

In 1997, Argentinean scientists Adrian Brunini and Mirta Parisi published a paper giving plausible ways that Uranus became tilted. They believed that if a collision had taken place, it had to be when Uranus was a more solid core surrounded by a planetary envelope. The impacting object would have hit the proto-Uranus from the opposite direction as it traveled around the Sun. The two scientists thought that studying Uranus's satellites was the key to figuring out its odd axial tilt. They concluded that either the satellites of Uranus were created by the collision itself, or no collision happened. Brunini and Parisi's study found that Uranus's satellite Prospero (S/1999 U3) set a number of constraints on any possible conclusion. Therefore, they believe, it is possible that a new theory of solar-system formation is needed to explain the tilt of Uranus. A number of scientists seem to be shifting toward the second explanation: that Uranus formed tilted on its side, and that the nebular theory for the formation of the solar system fails to explain how this could have happened. Researchers have been working on finding a simulation that solves this and other oddities of the solar system.

In 2006, Brunini published a new theory of the formation of the solar system in *Nature* magazine. His mathematical model is based on the idea that Jupiter and Saturn once had a 1:2 orbital resonance. This means that in the time it took Saturn to orbit the Sun once, Jupiter went around twice. The gravitational effect of Jupiter and Saturn gradually changed the orbits of Uranus and Neptune. Brunini's simulation shows that his model would take about a million years for the outer planets to reach the orbital positions we now observe. He argues that during the close encounter of Saturn and Uranus, the angular momentum of the planets shifted, which over time caused their axial tilts to change. This scenario, Brunini argues, can explain the orbits of Uranus's rings and satellites, which would have slowly changed their orientation along with Uranus. Unlike a collision, Brunini's scenario would have taken hundreds of thousands of years to play out.

No definitive answer has been found for what caused Uranus's unique tilt. Only Voyager 2 has visited Uranus; new spacecraft would be able to provide more data but cannot be sent until either the planets are again aligned for a "slingshot" (gravity-assist) approach (more than a century away) or the necessary nuclear propulsion systems are developed. Until then, researchers are left making mathematical and computer models in their efforts to solve the mysteries of Uranus's axial tilt.

METHODS OF STUDY

Uranus can be observed from Earth with telescopes, and on dark, clear nights can be viewed with the unaided eye. Scientists have also taken photographs of the Uranus system using the Hubble Space Telescope. In late 2002, astronomers in Chile were able to image Uranus, its rings, and some of its satellites. The pictures were taken with the Very Large Telescope (VLT) at the European Southern Observatory (ESO) Paranal Observatory. The rings that are normally unable to be viewed from Earth, along with seven satellites, appeared in the image because it was taken at near-infrared wavelengths.

The Voyager program is the only spacecraft that has visited Uranus. Launched in 1977, Voyager 2 came within 81,500 kilometers of Uranus on January 24, 1986. Voyager 2 was equipped with more than a dozen scientific instruments, including cameras, television cameras, magnetometer, and spectroscopes. Voyager 2 viewed Uranus's "south" pole (located south of the orbital plane), which was pointed toward the Sun. At Uranus, Voyager 2 discovered ten satellites and two rings. The spacecraft also studied the planet's five largest moons (Oberon, Umbriel, Titania, Ariel, and Miranda), taking the first close-up photographs of them.

Voyager 2 provided the first close-up photographs of Uranus and detailed information about its magnetic field, ring system, weather, and unusual axial tilt. By the early nineteenth century, scientists knew that the planet was tilted, but it was not until Voyager 2 arrived at Uranus that astronomers knew precisely how tilted it was.

Computer modeling can be used to explore the dynamics of complex systems over time, where geologic time is essentially replaced by computation time. Modern computers allow a tremendous amount of computational power, and the magnitude of that computational power is continuously increasing. Basically, a computer modeling effort such as that used by Brunini and similar researchers seeks to begin with certain basic assumptions about initial conditions of a complex system such as Uranus in its interaction with larger bodies such as Jupiter and Saturn, and then introduce the gravitational interactions between all of these bodies and allow the computational cycle to mimic the passage of time as each of these bodies orbits the Sun and continues to interact with the others. This sort of thing cannot be easily done by hand. Sir Isaac Newton, in presenting his development of mechanics in *Philosophiae Naturalis Principia Mathematica* (1687; commonly known as *The Principia*), provided a means of quantifying the gravitational interaction between two bodies. That relatively simple problem can be solved in closed form in both spatial and temporal coordinates. However, the three-body problem requires numerical analysis, which is largely done at present by computer programming or software packages, since it cannot be solved in closed form. The more bodies are involved in a calculation, the more computing power is required.

CONTEXT

Scientists may never know the real reason for Uranus's axial tilt. Further study of the planet Uranus by spacecraft, and even possibly by humans, could lead to the answer. Computer simulations and mathematical models can help scientists speculate what might have happened. Maybe the accepted nebular theory for how the solar system formed is incorrect. Maybe the gravitational effects of Jupiter and Saturn slowly caused Uranus to lean to its side. Maybe Adrian Brunini and his colleagues are correct, and the only way to make this determination is to study Uranus's satellites. The quest to explain Uranus's extreme axial tilt could lead to a new view of how Earth and the solar system formed.

Jennifer L. Campbell

FURTHER READING

Burgess, Eric. *Uranus and Neptune: The Distant Giants*. New York: Columbia University Press, 1988. Covers the Voyager 2 spacecraft's mission, technical difficulties, and its encounters with Jupiter, Saturn, and Uranus. Focuses on data collected by Voyager 2 about Uranus. Includes several illustrations and tables. Well written, suitable for the general audience.

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. A well-written college textbook for introductory astronomy courses. Includes a chapter on Uranus and Neptune.

Elkins-Tanton, Linda T. *Uranus, Neptune, Pluto, and the Outer Solar System*. New York: Chelsea House, 2006. Explores the Sun's relationship with the three outer planets and their moons. Looks at these planets as recorders of the formation of the solar system. Aimed at a general or high school audience. Illustrations, bibliography, index.

Encrenaz, Thérèse, et al., eds. *The Outer Planets and Their Moons: Comparative Studies of the Outer Planets Prior to the Exploration of the Saturn System by Cassini-Huygens*. New York: Springer, 2005. An in-depth look at the current understanding of the solar system's outer planets. Focuses on the studies of their formation, evolution, magnetospheres, satellites, and ring structures. For scientists, first-year graduate students, and advanced undergraduates.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. An introductory college text that gives students easy-to-understand analogies to help them with more complex theories.

Well written and easy to read. Includes a CD-ROM featuring InfoTrac software.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A thorough and well-written introductory college astronomy textbook. Covers all aspects of Uranus.

Hunt, Garry E., and Patrick Moore. *Atlas of Uranus*. New York: Cambridge University Press, 1988. This was the first volume after the 1986 Voyager encounter to offer a comprehensive history of Uranus: its discovery, satellites, rings, and the data returned by Voyager, including photographs.

Loewen, Nancy. *The Sideways Planet: Uranus*. Mankato, Minn.: Picture Window Books, 2008. An educational children's book devoted to the planet Uranus. Covers Uranus's rings, moons, and tilted axis.

Miner, Ellis. *Uranus: The Planet, Rings, and Satellites*. New York: Ellis Horwood, 1990. The author thoroughly covers the topics of both the Uranian system and the Voyager mission. Illustrations, bibliography, index.

Schmude, Richard W. *Uranus, Neptune, and Pluto and How to Observe Them*. New York: Springer, 2008. Ideal for backyard or amateur astronomers who are interested in observing the outer planets. Also includes up-to-date information about the planets.

Tocci, Salvadore. *A Look at Uranus*. New York: Franklin Watts, 2003. As part of the Out of This World series, this book covers all aspects of the planet Uranus, from its discovery through 2002. Includes several photographs. Suitable for all readers.

See also: Auroras; Earth's Atmosphere; Earth's Composition; Eclipses; Infrared Astronomy; Jovian Planets; Planetary Atmospheres; Planetary Formation; Planetary Interiors; Planetary Magnetospheres; Planetary Ring Systems; Planetary Rotation; Planetary Satellites; Planetology: Comparative; Telescopes: Ground-Based; Uranus's Atmosphere; Uranus's Interior; Uranus's Magnetic Field; Uranus's Rings; Uranus's Satellites.

V

Van Allen Radiation Belts

Category: Earth

The Van Allen radiation belts are concentrated rings of ionized particles in Earth's magnetosphere. Detailed study of the radiation belts led to an understanding of certain phenomena occurring in the ionosphere and the determination of the physical properties of the exosphere.

OVERVIEW

The Van Allen radiation belts are concentrated, torus-shaped regions of charged particles within Earth's magnetosphere. These particles, made up of protons, electrons, and other ions, spiral about in great numbers between Earth's magnetic poles. The magnetic and charged particles within the Van Allen belts can be divided into four regions: the Van Allen geomagnetically trapped radiation region, the auroral region, the magnetosheath, and interplanetary space. The inner and outer belts are part of the Van Allen geomagnetically trapped radiation region. In discussions of the Van Allen belts, the magnetic storm is often referred to as a third radiation belt.

The inner zone stretches from about 1,000 to more than 5,000 kilometers above Earth. It is mainly independent of time. Its composition is nearly consistent with that expected for the decay products of cosmic-ray-produced neutrons in the atmosphere (a neutron is an elementary, neutral particle of mass); this zone is of cosmic-ray origin. The radiation in the middle of the inner zone is composed of electrons with energies exceeding 40 kilo-electron volts (keV) and protons with energies greater than 40 million electron volts (MeV). (Electrons are elementary particles with a negative charge; protons are positively charged elementary particles.) In the inner belt, many of the high-energy protons are capable of penetrating several inches of lead. At the edge of the inner zone, in the region of geomagnetic latitudes 35° to 40°, low-energy electrons are found. The decay of light-scattering neutrons gives rise to high-energy protons. Beyond Earth's magnetic field, the mean ionizing capacity is 2.5 times higher than the minimum ionizing capacity. Particles in the inner zone are stable and exist for a long period of time.

The outer zone stretches about 15,000 to 25,000 kilometers above Earth. This zone undergoes very large temporal fluctuations appearing to be caused by solar activity and auroras, atmospheric heating, and magnetic storms. The outer belt contains soft particles; it is of solar origin. The outer zone contains electrons with more than 40 keV in energy and protons with more than 60 MeV in energy. The outer zone has greater geophysical significance than the inner zone. According to the comparison of E. V. Gorchakov, the boundaries of the outer zone coincide with isochasms (lines of equal probability of auroras). Trapped particles introduce magnetic effects in the outer radiation belt. This effect was measured by Luna 1. The increase in ionization of the outer zone is unstable. Particles exist for a short period of time compared with those of the inner belt.

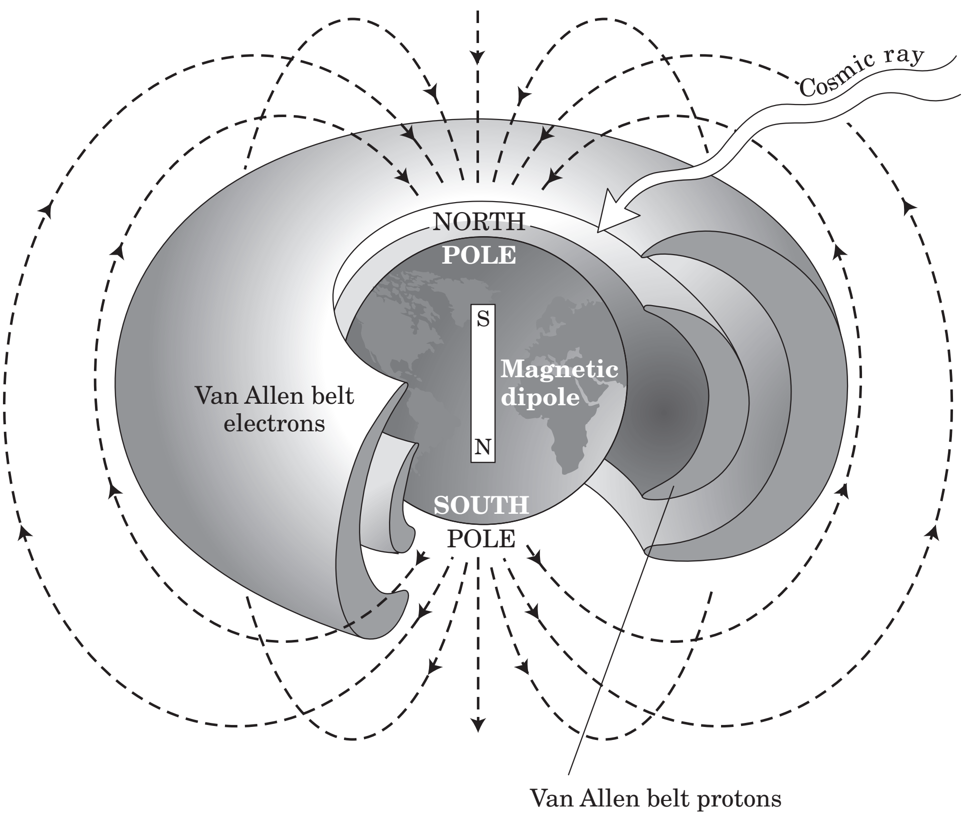
A third radiation belt is produced by magnetic storms. Protons are transported from the Sun in a corpuscular stream and injected by magnetic field perturbations into Earth's field. The charge exchange with neutral hydrogen in Earth's exosphere is the fastest mechanism of removal. This is about a hundred times faster than scattering from ions in the exosphere. With the exception of trapped radiation, the entire region in the magnetic cavity is known as the auroral region. Auroral particles, the islands or pulses in the long tail and spikes at high latitudes of 1,000 kilometers, are phenomena that occur in the auroral region. Electrons of uniform angular distribution have a roughly constant intensity between 100 and 180 kilometers in altitude.

The magnetosheath lies between the shock front formed by the solar wind and the magnetic cavity. Islands of electrons have been observed in the magnetosheath. At its widest, the magnetosheath is about four times the radius of Earth. It contains a compressed, seemingly chaotic interplanetary magnetic field. The interplanetary field connected to the Sun is predominantly in the ecliptic plane. The field terminates when the solar wind undergoes a shock transition to subsonic flow.

The lifetime of trapped particles decreases with distance from Earth. The lifetime of electrons with energies greater than 1 MeV at a distance of 1.2 to 1.5 times Earth's radius is about a year. The lifetime of the same electrons is reduced to days and months at a distance of 1.5 to 2.5 times Earth's radius. At even greater distances, the lifetime of the particles is measured in minutes. Because Earth is strongly influenced by the Sun's magnetic field, Earth's geomagnetic field does not decrease indefinitely with increasing distance. The solar wind pushes Earth's magnetic field and is deflected by it. At about 10 Earth radii, the radiation belt ends abruptly.

Particles of trapped radiation may be lost in two ways. During a magnetic storm, the magnetosphere may lose or gain particles. This occurs at distances of 1.0 to 1.5 times Earth's radius. The other mechanism occurs at distances greater than 8 times Earth's radius. Small, rapid variation in the magnetic field at such distances scatters trapped particles, dumping them into the atmosphere. In a similar fashion, it is seen that charged particles in Uranus's magnetosphere are swept down into the planet's upper atmosphere by collisions with particles in its ring system.

The Van Allen Radiation Belts



Beautiful auroral displays occur when the charged particles are dumped into Earth's upper atmosphere. Solar flares eject into space streams of high-energy protons and electrons. When these beams of high-energy particles are directed toward Earth, Earth's magnetic field is partially disrupted. Particles trapped within the field lines can escape downward toward Earth at the lower ends of the radiation belts. High-energy particles, reinforced with particles from the Sun, energize the upper atmosphere, causing luminous and often colorful auroras.

KNOWLEDGE GAINED

The years 1957 and 1958 were designated the International Geophysical Year (IGY), an international scientific tour de force to advance understanding of Earth sciences. As contributions to IGY, the Soviet Union and the United States both pledged to place a satellite in orbit about the Earth. Russia's Sputnik 1 beat the American effort to orbit. However, the American effort was the first to gather useful scientific information. With the data returned by Explorer 1, America's first artificial satellite, a high-energy radiation belt was detected by James A. Van Allen and his assistants, George H. Ludwig, Carl E. McIlwain, and Ernest C. Ray. The same observations were made by Explorer 3, launched by the U.S. Army on March 26, 1958, and Sputnik 3, launched by the Soviet Union on May 15, 1958.

Later, a satellite was launched as part of Project Argus, which studied the location, height, and yield of electron blasts. This project was carried out by the Advanced Research Projects Agency. The belt of electrons produced by the Argus nuclear explosions developed at a distance of twice Earth's radius. Explorer 4, launched on July 26, 1958, carried four Geiger counters to handle high levels of radiation. One of these Geiger counters was shielded with a thin layer of lead to keep out most of the radiation. The satellite reached a height of 2,200 kilometers and registered an intensity of high-energy radiation. From the data returned, scientists concluded that Earth is surrounded by belts of high-energy radiation consisting of particles originating from the Sun and trapped in the lines of force of Earth's magnetic field. These were named the "Van Allen radiation belts."

Explorer 4 obtained a kidney-shaped intensity contour of Earth's inner belt. Data from early Pioneer spacecraft suggested a solar origin of soft particles populating the outer zone. Three Pioneer probes and Luna 1 discovered the crescent-shaped intensity contours of the outer belt. Sputnik 3 data helped identify the bulk of the outer belt particles as low-energy electrons (10 to 50 keV).

Several more human-made belts were produced in 1962. The Starfish project, an American venture, created a belt much wider than the Argus belt. Decay of some of the particles took several years in low altitudes. In the same year, the Soviets created at least three similar belts. More sophisticated versions of the instrumentation used in these early probes of Earth were incorporated into spacecraft sent to other planets in the solar system. Probes to Mercury, Jupiter, Saturn, Uranus, and Neptune have discovered radiation belts similar to Earth's Van Allen belts. A planet needs a magnetic field to trap charged particles into a radiation belt or system of radiation belts. For that very reason, there are no significant belts of trapped charged particles at Venus or Mars.

CONTEXT

In the earliest days of space exploration, gauging the intensity of Earth's radiation belts with uncrewed spacecraft was crucial as a first step toward sending humans into space. Both the United States and the Soviet Union had a vested interest in the results of early investigations of the magnetosphere.

The Van Allen belts, while lifesaving in that they keep dangerous radiation from reaching the surface of the planet, are potentially hazardous to Earth-orbiting spacecraft. They threaten electronics systems and instrumentation and can interfere with radio transmissions. In the late 1950's, it was not known just how hazardous the radiation surrounding Earth would prove to humans. While it would not be wise to base a space station within the radiation belts, the belts themselves pose little threat to humans, who quickly punch through on voyages of exploration beyond the Earth; that was the case back in the Apollo program and will be true of the National Aeronautics and Space Administration's (NASA's) planned Constellation program flights to the Moon and later to Mars. However, leaving the safety of orbit beneath the Van Allen belts does expose astronauts to the potential hazards of ionizing radiation streaming outward into the solar system from the Sun. Therefore, special protection must be provided to humans on expeditions beyond low-Earth orbit.

The relationship between auroras and the Van Allen belts has been studied for decades, but although the overall phenomenon has been well characterized, all is not completely understood. Scientists do know that most bright auroras are produced by electrons dumped into Earth's atmosphere by solar flares. The auroral particles are the electrons escaping from the outer Van Allen radiation belt. The average kinetic energy of the electrons is 32 keV. The leakage of corpuscular radiation into the auroral zones is the most important loss of corpuscular radiation from the outer Van Allen belt.

Satya Pal

FURTHER READING

Bone, Neil. *The Aurora: Sun-Earth Interaction*. New York: John Wiley, 1996. One volume in the Ellis Horwood Library of Space Science and Space Technology. Devoted to describing the electrodynamics of the Sun-Earth environment that produce auroral displays.

Bothmer, Volker, and Ioannis A. Daglis. *Space Weather: Physics and Effects*. New York: Springer Praxis, 2006. A selection from the publisher's excellent Environmental Sciences series, this is an overview of the Sun-Earth relationship and provides a historical and technological survey of the subject. Projects the future of space weather research through 2015 and includes contemporary spacecraft information.

Foerstner, Abigail. *James Van Allen: The First Eight Billion Miles*. Iowa City: University of Iowa Press, 2007. An engaging portrait of the legendary physicist, discussing his contributions to the World War II effort as well as to the advancement of studies of Earth's geomagnetic environment, his early efforts to study cosmic rays using balloon-launched rockets, the Explorer 1 story, and Van Allen's continuing participation in studying space physics until his passing in 2006.

Gregory, Stephen A. *Introductory Astronomy and Astrophysics*. 4th ed. San Francisco: Brooks/Cole, 1997. Suitable as a textbook for introductory college courses or advanced high school courses in general astronomy. Covers all topics from solar-system bodies to cosmology. Some errors and issues with mathematical presentations.

Kallmann-Bijl, Hildegaard, ed. *Space Research: Proceedings of the First International Space Science Symposium*. New York: Interscience, 1960. A detailed explanation of the theory behind the nature, origin, and composition of the inner, outer, and third radiation belts is provided. Dated and technical, but provides a useful historical perspective.

Milone, Eugene F., and William Wilson. *Solar System Astrophysics: Background Science on the Inner Solar System*. New York: Springer, 2008. Rigorous and highly mathematical presentation involving geophysics, atmospheric physics, and mineralogy covering all aspects of planetary science. Includes results from the Mars Exploration Rovers and Cassini spacecraft.

Moldwin, Mark. *An Introduction to Space Weather*. Cambridge, England: Cambridge University Press, 2008. This text introduces space weather, the influence the Sun has on Earth's space environment, to the nonscience reader. Discusses both the scientific aspects of space weather and issues of technological and societal import.

Savage, Candace. *Aurora: The Mysterious Northern Lights*. New York: Firefly Books, 2001. Heavily illustrated with photographs of auroral displays, this book provides a history of scientific investigation of auroral phenomena.

Sullivan, Walter. *Assault on the Unknown*. New York: McGraw-Hill, 1961. Sullivan describes the discovery of the Van Allen belts in great detail.

Van Allen, James A. *The Magnetospheres of Eight Planets and the Moon*. Oslo: Norwegian Academy of Science and Letters, 1990. A technical summary of all major magnetic structures in the solar system, written by the prolific researcher after whom the Van Allen belts are named.

_______. "Radiation Belts Around the Earth." *Scientific American* 200 (March, 1959): 39-47. A seminal article by the discoverer of the Van Allen belts. Well illustrated. Provides insight into the nature of the original discovery and the early days of the space race.

See also: Auroras; Earth's Atmosphere; Earth's Magnetic Field: Origins; Earth's Magnetic Field: Secular Variation; Earth's Magnetic Field at Present; Earth's Magnetosphere; Earth-Sun Relations; Jupiter's Magnetic Field and Radiation Belts; Planetary Magnetospheres; Solar Flares; Solar Wind.

Venus's Atmosphere

Categories: Planets and Planetology; Venus

The atmosphere of Venus has a surface temperature of about 743 kelvins and a surface pressure of about 90 Earth atmospheres. Its clouds consist largely of carbon dioxide, and droplets of sulfuric acid rain down to the surface.

OVERVIEW

The second planet from the Sun and Earth's immediate inner neighbor, Venus is often called Earth's twin, because the masses and radii of the two planets are very similar. Venus's mass is 82 percent that of Earth. Its radius is only 5 percent less than Earth's. Under ordinary circumstances, no objects in the sky other than the Sun and the Moon surpass Venus in brightness. Viewing Venus with one or more of the other planets also visible in the sky shortly before sunrise or briefly after sunset can be an awe-inspiring sight. In ancient times, since Venus could at times be seen in the morning skies and at other times in the evening skies, the planet was actually thought to be two different objects; they were given the names Phosphoros and Hesperus for the morning and evening star, respectively.

Venus revolves once around the Sun every Earth 224.7 days. It rotates on its axis once every 243.01 days. The inclination of its orbit is about 3.5° with respect to the ecliptic plane. Its orbit, like those of all the other planets, is an ellipse, but it is very close to being a perfect circle. The tilt of Venus's axis of rotation is about 17.8°, as compared with the Earth's 23.5°. As a result, any seasonal changes of weather on Venus would be less extreme than on Earth. Remarkably, its rotation is retrograde, backward, as compared with the direction of revolution, or clockwise, as seen from above the ecliptic plane. It is not known with certainty why Venus has a retrograde rotation. Virtually all other objects in the solar system rotate and revolve prograde, or counterclockwise as seen from above the ecliptic.

At each inferior conjunction (each time Venus and the Sun are aligned in the sky with Venus closer to Earth than to the Sun), the same face of Venus points toward the Earth. This phenomenon may mean that Venus's rotation is influenced by the Earth's gravitational pull; however, some information indicates that the alignment at inferior conjunction is not exactly perfect and therefore may be coincidental.

As seen from Earth, Venus's maximum disk size, or angular diameter on the sky, is about 0.02°, and its minimum angular diameter is about 0.003°. The maximum or minimum size on the sky corresponds to the closest and farthest distances from Earth. In contrast, the Sun and the Moon have an apparent angular diameter of about 0.5°. Given these observational circumstances, little can be learned about the planet's atmosphere from telescopic observations. The planet's average density is 5.25 grams per cubic centimeter, and the surface acceleration caused by gravity is 0.903 times that of Earth. Escape velocity from Venus's surface is 10.3 kilometers per second. That means Venus can retain its atmosphere virtually indefinitely, since most of the molecules at the top of the atmosphere will travel at speeds well under 10.3 kilometers per second.

Through a telescope, an image of Venus is somewhat disappointing; only the planet's phases are obvious. Exceptional atmospheric and viewing conditions at good astronomical

This image of one complete Venusian hemisphere is a composite of high-resolution Magellan images whose gaps are filled by images from the Arecibo radio telescope in Peru. (NASA/JPL/USGS)

sites have allowed scientists to see and photograph subtle variations in shading on bright cloud tops of the planet's lit side. However, no part of the surface of Venus is visible through telescopes, because of its thick cloud cover.

Sulfur is probably released into the atmosphere of Venus by outgassing volcanic processes. Sulfur rises and bonds with water and oxygen to produce sulfuric acid. It appears as a fairly thick haze and is more strongly concentrated than the acid found in the battery of an automobile. These clouds are less murky than fog, with visibilities within them of perhaps several hundred meters. Top layers of these clouds are about 80 kilometers above the planet's surface. The tops of the sulfuric acid-laden clouds have winds that move faster than 300 kilometers per hour, a speed comparable to that of the Earth's jet streams. These high winds swirl around the planet in about four days. Circulating motions cause gas to rise near the equator and descend near the poles, probably a direct consequence of excess solar heating in the equatorial region.

The atmosphere of Venus is mostly carbon dioxide. Venus has undergone a spectacular greenhouse effect, caused by particles in the clouds trapping or absorbing infrared radiation. Most sunlight is reflected by clouds back into space. The albedo of Venus is about 0.8; that is, 80 percent of incident radiation striking the cloud tops is reflected back into space. The 20 percent that penetrates the clouds warms the surface sufficiently to heat the surface rocks and terrain. These surface structures radiate, essentially at infrared wavelengths, and that radiation cannot penetrate the clouds and gets trapped. Heat builds, and the high temperature releases gases (mainly carbon dioxide) from rocky minerals into the atmosphere, which in turn drives the greenhouse effect further. A cycle develops: The carbon dioxide traps the radiation, which heats the surface and atmosphere, which then triggers the release of more carbon dioxide. This continues until the entire atmosphere reaches a sufficiently high temperature (740 kelvins) to radiate as much energy back into space as it receives.

High above the clouds, the atmospheric layers called the exosphere consist mostly of hydrogen and helium. These strata are affected by intense incoming ultraviolet solar radiation. The radiation ionizes atoms, making them electrically charged. The uppermost layers form the ionosphere. Venus's ionosphere is not as intensely ionized as Earth's.

Venus's surface is quite flat compared with that of the Earth; however, there are at least three major elevated regions, and they exhibit features that influence the atmosphere. The largest, Ishtar Terra, is similar in size to Australia. The others are about the size of the largest islands in Indonesia. Many less elevated regions are also present. On Ishtar Terra, a mountain called Maxwell Montes is apparently higher than Mount Everest with respect to the surrounding flat terrain. Large volcanoes are present on each of the three plateaus. Volcanoes imply gas release from the planetary interior; therefore, Venus's atmosphere is attributable in part to volcanic outgassing. Although the atmosphere is composed mostly of carbon dioxide, there is some nitrogen and sulfur dioxide, very little water vapor, and trace amounts of various other gases. Sulfur dioxide is outgassed by volcanoes on the Earth but is quickly diluted by rain and moisture. In contrast, sulfur dioxide outgassed in the dry atmosphere of Venus is very stable, accounting for the efficient production of sulfuric acid in the clouds and elsewhere.

Continent-building processes caused by plate tectonics on the Earth may have occurred on Venus, though to a far lesser extent. This idea is based mostly on the fact that only a few elevated regions exist. Therefore, the outgassing brought on by a variety of volcanic actions related to plate tectonics is probably slower and less effective on Venus, compared with the heat-induced gases, such as carbon dioxide, from rocks.

One other feature of Earth's atmosphere that probably does not exist to any extent on Venus is production of high-level auroras (or, as they are known on Earth, the northern and southern lights), because Venus has little or no magnetic field. Since its rotation rate is very slow, one would expect a weak but nevertheless measurable magnetic field. Several theories have been put forth to explain this lack. One theory is that, like the Earth, Venus undergoes polarity changes of its overall dipolar magnetic field. The Earth's magnetic field is explained by the dynamo hypothesis. The rotating core produces loop currents in the heated molten regions, which in turn produce a magnetic field. Geological and paleontological evidence suggests that the Earth's magnetic field undergoes reversals of polarity at irregular intervals. During the reversal periods, little or no field is present. It is possible that Venus could be undergoing such a magnetic reversal phase. In fact, either it is in such a phase or the dynamo hypothesis is incorrect.

Venus's Atmosphere Compared with Earth's

	Venus	*Earth*
Surface pressure (bars)	92	1.014
Surface density (kg/m^3)	~65	1.217
Avg. temperature (kelvin)	737	288
Scale height (kilometers)	15.9	8.5
Wind speeds (meters/second)	0.3-1.0	up to 100
Composition		
Argon	70 ppm	9,430 ppm
Carbon dioxide	96.5%	350 ppm
Carbon monoxide	17 ppm	—
Helium	12 ppm	5.24 ppm
Hydrogen	—	0.55 ppm
Hydrogen chloride	tr	—
Hydrogen fluoride	tr	—
Krypton	—	1.14 ppm
Neon	7 ppm	18.18 ppm
Nitrogen	3.5%	78.084%
Oxygen	—	20.946
Sulfur dioxide	150 ppm	—
Water	20 ppm	1%
Xenon	—	0.08 ppm

Note: Composition: % = percent; ppm = parts per million; tr = trace amounts.

Source: Data are from the National Space Science Data Center, NASA/Goddard Space Flight Center.

METHODS OF STUDY

The first attempt to send an interplanetary probe to Venus did not fare well. Mariner 1 launched on July 22, 1962. A software error caused its booster to veer dangerously off course while low in the atmosphere, and it was destroyed on purpose by the range safety personnel at Cape Canaveral.

Mariner 2, a sister spacecraft to the failed first American Venus probe, launched successfully on August 27, 1962. Fortunately, this probe was able to fly within

34,833 kilometers of the Venusian surface on December 14, 1962. Although it carried no photographic equipment, Mariner 2 provided a treasure trove of new information about the shrouded planet. The spacecraft was outfitted with Geiger tubes, an ion chamber, a cosmic dust detector, a microwave radiometer, and a magnetometer experiment. Mariner 2 determined that the Venusian surface temperature was more than 670 kelvins. It detected neither a planetary magnetic field nor any Van Allen-like radiation belts about the planet. It continued to collect data about particles and fields in interplanetary space until contact was lost on January 3, 1963, at a distance of 87 million kilometers from Earth.

Mariner 5 was launched on June 14, 1967, and was sent to fly by Venus. This spacecraft also did not include photographic or television cameras. It was equipped with radio science and ultraviolet experiments as well as particle and magnetic field detectors. Mariner 5 encountered Venus on October 19, 1967, coming within 4,000 kilometers. The spacecraft investigated Venus's cloud tops and the solar wind interacting with interplanetary magnetic fields.

Early images taken by the space probes Mariner 10 and Pioneer Venus in reflected solar ultraviolet light revealed considerable variation in shading on Venus. Variation is caused by radiation coming from different levels of the clouds. The study of the motion of these clouds has indicated that the top layers can rotate at very high speeds, approaching perhaps 100 kilometers per hour. These high-strata, rapid wind velocities are in part caused by the hot, sunlit clouds transferring heat to the colder dark side. The cloud structure shows three distinct strata: a high, thick layer; a medium-high haze layer; and a lower, medium-thick layer. From this lower layer downward it is essentially clear all the way to the surface.

The Pioneer Venus atmospheric probes, sent to the surface via a combination of small retrorockets and parachutes, and Soviet Venera landers measured a decrease in the wind velocity at lower altitudes. On the surface, winds are essentially gentle breezes. Heat transfer and exchange in the atmosphere are very dependent on the density and pressure of various layers. Lower levels of the atmosphere are under superhigh pressures. Soviet Venera landers and the Pioneer space probes measured pressures of about 90 atmospheres at the surface. Heat transfer is so efficient that there is no large-scale difference between daytime and nighttime temperatures at the surface. Mariner 10 and Pioneer Venus ultraviolet images indicated an overall circulation pattern: Atmospheric gas rises at the equator and descends at the poles. With slow rotation of the planet, this circulation pattern is highly stable.

Four Venera landers managed to set down on the surface, perform experiments, and obtain electronic images of the surroundings using a fish-eye lens or wormlike view of the terrain to the horizon. At both landing sites, rocks and the horizon in the clear atmosphere are visible. Rocks are clearly of volcanic origin. In some cases their sharp edges indicate little or no erosion, suggesting recent volcanic origin. One would expect that erosion, under the high pressure and intense heat, would quickly deform and erode the rocks' edges. Pioneer Venus detected the existence of sulfuric acid droplets, which at the high temperatures and pressures present is very corrosive.

Pioneer Venus was equipped with a radar ranger. Scientists could send radar beams to the surface of Venus from the spacecraft and measure the time interval from the emission of the radar to the subsequent receiving of the reflected echo. This procedure allowed accurate measurement of distances between the spacecraft and the surface. After a compilation of such observations, the Pioneer Venus mission team was able to provide a detailed map of the surface terrain for the first time. Better maps would have to await a more sophisticated radar system placed in orbit about Venus.

Vega 1 and 2 were ambitious missions involving identical carrier spacecraft that each delivered both a lander based on the Venera design and an instrumented balloon to Venus before both carrier spacecraft were then redirected to join an international group of spacecraft intercepting and studying Halley's comet near its 1985/1986 perihelion passage. The carrier craft were outfitted with an imaging system, an infrared spectrometer, and a spectrometer capable of ultraviolet through infrared observations, detectors of dust and micrometeoroids, a plasma energy analyzer, a magnetometer, wave and plasma analyzers, a neutral gas mass spectrometer, and an energetic particle analyzer. The Vega 1 and 2 carrier craft encountered Venus on June 11 and 15, 1985, respectively, having several days earlier ejected their lander and balloon payloads. Neither carrier craft provided deep new insights into the nature of the Venusian atmosphere, but they did set the stage for the unique balloon payload and their results.

Venus's atmosphere apparently provided a particularly strong wind gust when the Vega 1 lander was still 20 kilometers above the planet's surface. This even activated the surface experiments early, and no results were produced after touchdown. Vega 2 operated properly and on June 15, 1985, safely touched down in the eastern Aphrodite Terra region. This lander determined the local atmospheric pressure to be 91 atmospheres, with a surface temperature of 736 kelvins. It endured the extreme environment for just under an hour, but, before it failed, the lander determined a rock sample to be a variety of anorthesite.

The Vega 1 and 2 balloons floated in the planet's atmosphere, providing data for about forty-six hours at an altitude of approximately 54 kilometers. The balloons were small in mass and size (25 kilograms, about 55 pounds, and 3.4 meters, or 11 feet, in diameter) but were able to dangle a gondola assembly filled with instruments to sample and measure the Venusian atmosphere. The balloons began their mission after being deposited on Venus's dark side. They sank to a depth of 50 kilometers before rising again to an altitude of 54 kilometers where they determined the pressure and temperature to be similar to those conditions on Earth. However, wind speed was nearly that of hurricane status, and at this altitude the carbon dioxide-rich atmosphere had a strong concentration of sulfuric acid with far less hydrofluoric and hydrochloric acid. Before losing electrical power, the balloons registered a variable vertical component to the atmospheric winds upon which they floated. Also they survived long enough to move from the dark side to the planet's illuminated side.

The National Aeronautics and Space Administration's (NASA's) next probe to Venus was named after the great Portuguese explorer Ferdinand Magellan. Its goal was no less daunting than to use an imaging radar to map at least 98 percent of Venus's surface to a resolution of 100 meters or less. Magellan was deployed from the space shuttle *Atlantis* on the STS-30 mission on May 5, 1989, and sent on its way toward the inner solar system. The spacecraft arrived in Venus orbit on August 10, 1990. Its synthetic aperture radar system was able to peer through the thick atmosphere. After four years of mapping, radar altimetry, and gravitational field measurements, NASA intentionally drove Magellan through the planet's atmosphere on October 12, 1994. In one final experiment, information was inferred about the atmosphere as the spacecraft heated up, and contact was eventually lost when Magellan was destroyed.

The European Space Agency (ESA) launched the Venus Express spacecraft on November 9, 2005. Largely a twin of ESA's successful Mars Express spacecraft, but modified to study Venus, Venus Express entered a nine-day-period polar orbit about Venus on April 11, 2006. Then for science operations to commence, that orbit was altered to have a twenty-four-hour period. Equipped with a penetrating radar, Venus Express began generating a surface map of the planet at resolutions even better than Magellan had achieved. However, science objectives of Venus Express also included detailed atmospheric studies. Aboard the spacecraft were infrared, visible-spectrum, and ultraviolet instruments to observe Venusian atmospheric characteristics and determine temperature profiles as a function of altitude.

Venus Express discovered a rather unexpected double vortex feature located around the south pole of the shrouded planet. This remarkable find occurred on the spacecraft's very first highly elongated orbit about Venus. Thus Venus Express was able to examine the planet's atmospheric patterns in ultraviolet and infrared from a global perspective (when far from Venus) and at close range (as it approached its low point in orbit). A vortex feature had been seen previously over the planet's north pole by earlier spacecraft, but a double vortex with a stable structure was quite unusual. Invoking the high wind speed of the upper atmosphere and the convection of rising hot air was insufficient to explain this double vortex. Using infrared sensors, Venus Express was able to map out windows in the atmosphere through which thermal radiation could escape to space. That modeling assisted scientists in determining cloud structures as a function of altitude above the tremendously hot planetary surface.

Some earlier probes had provided circumstantial evidence that lightning was present in Venus's atmosphere, but others produced data strongly suggesting there was a total lack of lightning. In 2006, the magnetometer aboard Venus Express provided definitive data that lightning does occur in Venus's atmosphere.

Venus Express made another important discovery in 2008, detecting hydroxyl molecules. This was the first time on a planet other than Earth that hydroxyl molecules had been detected; hydroxyl is a molecular ion consisting of one oxygen and one hydrogen atom bonded together covalently. Hydroxyl was detected at an altitude of 100 kilometers above the Venusian surface, which Venus Express accomplished by means of the Visible and Infrared Thermal Imaging Spectrometer, picking up the faint infrared light emitted by these molecules in a very narrow band of Venus's atmosphere. That band appears to be only 10 kilometers thick.

Hydroxyl has been found around comets. However, planetary atmospheres produce the molecule in a very different manner from that involved in comets. Hydroxyl on Earth is associated with the abundance of ozone in the upper atmosphere. Thus, detection of hydroxyl in Venus's atmosphere suggests that Venus still retains some Earth-like aspects. Absorption of ultraviolet light by hydroxyl molecules is important to the heating balance of any planetary atmosphere. The hydroxyl data from Venus Express would greatly assist planetary scientists in fine-tuning their models of the Venusian atmosphere.

Finally, in late 2008 Venus Express for the first time detected water being lost from Venus's daylight side. The previous year, this spacecraft's Analyzer of Space Plasma and Energetic Atoms detected the signature of hydrogen being stripped away from the planet's nightside. The orbiter's magnetometer was used to find hydrogen dissociated from water coming off the daylight side to be lost into space. Solar wind particles penetrate Venus's atmosphere, since the planet lacks a protective magnetic field. Scientists believe that solar wind particles break water molecules into two parts hydrogen and one part oxygen. Oxygen and hydrogen have been found escaping the nightside in the right proportion, but oxygen escaping from the daylight side was not seen in the 1:2 ratio required if the hydrogen seen comes from water. In any event, the solar wind mechanism was believed by many to be the means whereby, over time, Venus lost much of its original water.

CONTEXT

Earth's atmosphere has the potential to become more like that of Venus. There are two ways that such a situation could develop. If Earth moved closer to the Sun, increased solar heat would release more carbon dioxide into the atmosphere. Limestone rock (calcium carbonate) and dissolved carbon dioxide in the oceans provide a tremendous store of trapped carbon dioxide. Under higher temperatures, the rocks and seashells would chemically release carbon dioxide, and the greenhouse mechanism would raise the atmosphere's temperature. As a result, new carbon dioxide would be released and would speed the greenhouse mechanism. The second way that Earth's atmosphere could become more like that of Venus involves pollution of the atmosphere to the extent that enough carbon dioxide accelerates the existing greenhouse effect.

More generally, study of Venus's atmosphere helps scientists to better understand terrestrial weather and climate. Earth has an atmosphere composed mostly of nitrogen. The weather, however, is influenced primarily by water molecules and carbon dioxide molecules. These substances are found in Earth's atmosphere only in trace amounts; nevertheless, they are responsible for most of the heat transfer around the globe. On Venus, the weather is controlled by the atmosphere's main constituent, carbon dioxide. The contrasts between weather processes on Venus and those on the Earth have led to a more complete understanding of the latter. In an even larger context, comparative planetology studies of Venus, Earth, and Mars contribute to a better understanding of Earth's complex weather system and atmospheric physics. Moreover, such study helps us learn why three planets, all within the Sun's habitable zone, could evolve so differently. To understand Earth's evolution fully, it is necessary to know why the Venusian atmosphere became thick in carbon dioxide at great pressure (so that a planetary greenhouse effect led to runaway temperatures), and also to understand why the Martian atmosphere became thin in carbon dioxide at low pressure and low temperature. Earth, on the other hand, developed a nitrogen-oxygen atmosphere with traces of carbon dioxide; this led to reasonable temperatures and pressures and the development of a complex biosphere and interactive oceanic-atmospheric processes to maintain a dynamic equilibrium.

James C. LoPresto and David G. Fisher

FURTHER READING

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999. Amply illustrated with color images, diagrams, and informative tables, this book is aimed at a popular audience, but it can also be useful to specialists. Contains an appendix with planetary data tables, a bibliography for each chapter, planetary maps, and an index.

Cattermole, Peter John. *Venus: The Geological Story*. Baltimore: Johns Hopkins University Press, 1996. Provides a comprehensive presentation of the latest understanding of Venus based on Magellan data.

Elkins-Tanton, Linda T. *The Sun, Mercury, and Venus*. New York: Chelsea House, 2006. Examines the innermost portion of the solar system and the star, our Sun, which plays such a prominent role in the evolution of both planets. For the general audience with an interest in science.

Esposito, Larry W., Ellen R. Stofan, and Thomas E. Cravens, eds. *Exploring Venus as a Terrestrial Planet*. New York: American Geophysical Union, 2007. A collection of articles covering all major areas of planetary research on Venus. Technical.

Fimmel, Richard O., Lawrence Colin, and Eric Burgess. *Pioneering Venus: A Planet Unveiled*. Washington, D.C.: National Aeronautics and Space Administration, 1995. A complete summary of the findings of Pioneer Venus as of 1983. Pioneer Venus orbited Venus and sent several probes into the atmosphere of the planet. It also mapped the planet using a radar-ranging device.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. A college text on astronomy, somewhat more advanced than many introductory texts, but with a wealth of detail and excellent diagrams. Comes with a CD-ROM. The chapter on Venus is lucid and filled with spectacular diagrams and photographs.

Grinspoon, David Harry. *Venus Revealed: A New Look Below the Clouds of Our Mysterious Twin Planet*. New York: Basic Books, 1998. A thorough examination of the geology of Venus, incorporating Magellan mapping and other data. Explains the Venusian greenhouse effect. A must for the planetary science enthusiast who wishes to read an integrated approach to science and history. Includes speculation about Venus's past.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all aspects of planetary science. A comparative planetology approach is used rather than presenting just one chapter on all characteristics of Venus.

Marov, Mikhail Ya, and David Grinspoon. *The Planet Venus*. New Haven, Conn.: Yale University Press, 1998. Marov was Soviet Venera mission chief scientist, Grinspoon a NASA-funded scientist studying Venus. Together they provide a coordinated description of American and Soviet attempts to learn the secrets of Venus, a planet shrouded in mystery. For both general readers and specialists.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. A textbook at the beginning college level, introducing the scientific knowledge of the solar system as of 1988. The chapter on Venus goes into detail about the greenhouse effect and the contrasting atmospheres of Venus and the Earth.

Spanenburg, Ray, and Kit Moser. *A Look at Venus*. New York: Franklin Watts, 2002. A look beneath the thick clouds of Venus written for a younger audience.

See also: Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Earth-Sun Relations; Extraterrestrial Life in the Solar System; Greenhouse Effect; Mars's Atmosphere; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Rotation; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets; Venus's Craters; Venus's Surface Experiments; Venus's Surface Features; Venus's Volcanoes.

Venus's Craters

Categories: Planets and Planetology; Venus

Impact craters are the most numerous and most easily recognized surface features in the solar system. Because of their pristine nature, Venusian impact craters provide a unique opportunity for astrogeologists to study the effects of atmospheric variabilities and gravity in the formation of planetary surfaces.

OVERVIEW

One of the earliest and most crucial phases in the early development of the solar system was the Great Bombardment. During this phase, planetary surfaces were under intense bombardment by cosmic debris. For the Earth and Moon, the Great Bombardment began about 4.6 billion years ago and declined about 3. 8 billion years ago. Because the Moon has no active tectonism or atmospheric processes, impact craters there are mostly scars left from this time of intense bombardment. Because Earth, by contrast, is tectonically active and continuously subjected to weathering processes, craters of such ancient ages are not visible on Earth's surface and have been identified only on the cratons of continents, which undergo relatively little resurfacing from weathering, tectonism, volcanism, or other processes. The majority of impact craters identified on Earth, therefore, are geologically young and do not represent impacts from the Great Bombardment. Venus, Earth's closest planetary neighbor, presumably was also subjected to this intense cosmic bombardment, but it appears that Venus totally lacks the ancient, heavily cratered surface that occurs on Mercury, Mars, the Moon, and the rocky moons of other planets.

The detailed morphology of the surface of Venus is known through research done on remote-sensing data obtained from successful missions mounted by the National Aeronautics and Space Administration (NASA)—the Mariner (1962-1975), Pioneer (1978), Galileo (1990), Magellan (1990-1994), Cassini-Huygens (1998-1999), and MESSENGER (2004) uncrewed spacecraft missions—as well as the Soviet Union's Vega (1985) and Venera missions (1967-1984) and the European Space Agency (ESA) Venus Express (2005) mission. Other data have been accumulated through study of high-resolution images from Earth-based radar. The Magellan spacecraft, inserted into Venus orbit in 1990, yielded radar images with a resolution of a few hundred meters covering nearly 98 percent of the planet. In addition, the Soviet Venera missions mapped nearly 25 percent of Venus with additional radar imaging. These data have resulted in the mapping of nearly 100 percent of Venus and suggest that Venus lacks the ancient, heavily cratered surface occurring on other terrestrial planets and rocky satellites. The Venus Express mission has provided valuable data on the atmosphere of Venus and evidence suggesting that past oceans may have existed on the Venusian surface. A new Venus mission, Planet-C is planned for 2010 by the Japan Aerospace Exploration Agency (JAXA).

Astrogeologists use the density of impact craters to determine the age of planetary surfaces. The older the surface, the more impact craters it will have accumulated over time. On Venus this dating technique is problematic, because there are relatively few impact craters. Based on the density of Venus's impact craters larger than 30 kilometers in diameter, esti-

Venus's Dickinson Crater, 69 kilometers in diameter, from the Magellan spacecraft. (NASA/JPL)

mates of cratering rates scaled for other terrestrial planets and rocky satellites, and the known population of asteroids crossing Venus's orbit, the planet's average surface age is estimated at between 450 million and 250 million years old, the younger age being more likely. This suggests the surface terrain of Venus may be less than 5 percent of the age of the solar system. However, these ages are average estimates, and based on superposition some Venusian impact craters, volcanic structures, and tectonic terrains are thought to be as young as 50 million years old. Average age estimates aside, for the last 700 million years Venus has been subjected to significant surface volcanism, which probably is the reason that so few impact craters are visible on the planet's surface: Older impact craters have been covered over with lava flows or destroyed during episodes of catastrophic volcanic eruptions.

Venus has the densest atmosphere of any terrestrial planet. The Venusian surface pressure is equivalent to 94 bars—more than ninety times the pressure humans feel from Earth's atmosphere (90 bars is approximately the weight of water at 1 kilometer below the surface of Earth's oceans). In addition, Venus's atmosphere is composed of 96 percent carbon dioxide and trace amounts of nitrogen, water vapor, argon, carbon monoxide, and other gases. These clouds seal in the Sun's heat, creating a perpetual greenhouse effect that boosts surface temperatures on Venus to around 753 kelvins. Clouds within the Venusian atmosphere are composed mainly of sulfuric acid and small amounts of hydrochloric and hydrofluoric acid. The presence of such a dense atmosphere effectively filters the numbers of potential impactors by severely decreasing their kinetic energy during transit and preventing all but the largest incoming objects from impacting the Venusian surface. Craters smaller than 1.5 kilometers appear not to exist on Venus. Many craters of this size are distinctly noncircular and form groups or clusters of craters. This phenomenon is attributed to the impactor's becoming fragmented as it passes through the dense Venusian atmosphere and hitting the surface like a shotgun blast rather than like an artillery shell.

The variety of morphologies seen in Venus's impact craters tends to depend on their size. As the diameter of Venusian craters increases, changes in crater morphology take place and appear to correlate directly with Venus's surface gravity and dense atmosphere. Much of the morphology of Venusian impact craters is unique. Craters larger than 11 kilometers in diameter exhibit morphological characteristics similar to comparable complex craters on other planets: a circular shape, surrounding ejecta blankets, well-defined rims, terraced walls, central ring structures or central peak complexes, and, in the largest craters, multiple-ring basins. However, smaller Venusian craters tend to display a wide variation in shape and structural complexity—the opposite of the cratering patterns seen on other terrestrial planets and rocky satellites.

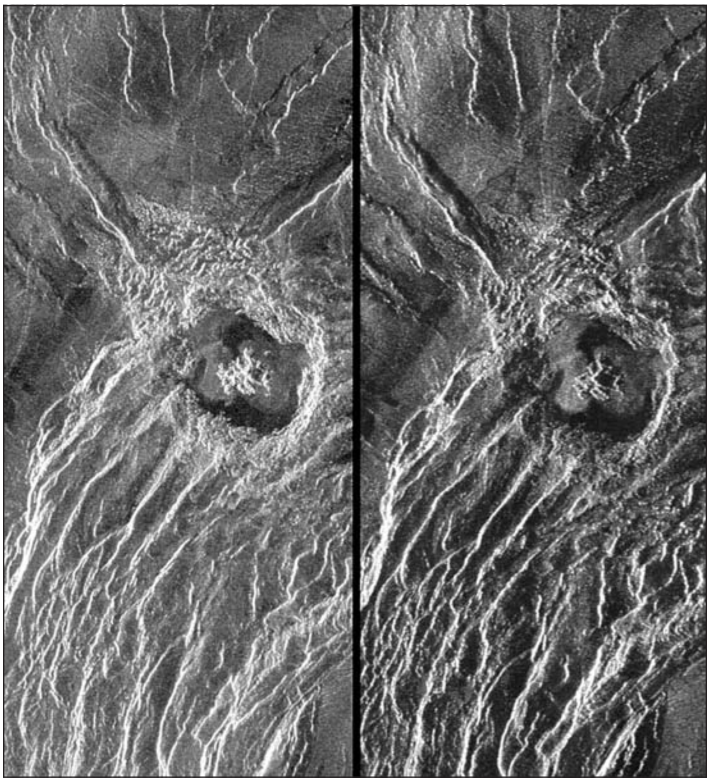
The morphological divergence is most directly attributed to the greater atmospheric density of Venus. Large multiring basins on Venus display at least two, and sometimes three or more, rings and near-pristine morphology; are surrounded by blocky ejecta distributed in lobes or raylike patterns; and in some cases produce lavalike flows of ejecta traveling several radii from the crater. The ejecta patterns are attributed to the dense Venusian atmosphere's slowing the travel path and speed of debris exiting the crater during impact. Many of the largest Venusian impact craters appear to have little to no topographic relief. Shallowness of these craters may be linked to Venus's lower gravity, producing slower impact speeds, and the planet's high surface and crustal temperatures, producing a large volume of impact-generated melt that remains in a near-molten state, allowing it to flow over long periods of time and eventually fill in the crater.

It is also suggested large Venusian impacts could trigger the subsequent volcanic or tectonic activity that disguises, or eventually erases, them within the landscape. One of the more difficult aspects of studying Venusian craters is distinguishing impact craters from circular volcanic calderas. High-resolution radar images help in defining the morphology of impact craters versus volcanic features by distinguishing ejecta deposits from lava flows. Unfortunately, lavas generated from impact-triggered volcanism can complicate discerning these structures because the lava flows may infill the impact craters, making them look like calderas.

One of the most unusual phenomena associated with Venusian impact craters is parabolic halos. These halos surround about 10 percent of the youngest craters and usually expand westward. The halos are attributed to the formation of a pre-impact bowshock wave created by the impactor's producing strong turbulence as it travels through the dense Venusian atmosphere. The turbulence lifts surface dust high into the air, and then prevailing easterly winds resettle the dust after the impact. Because the halos appear unaffected by volcanic, tectonic, or atmospheric processes, the haloed craters may be no more than 50 million years old, making them useful dating horizons.

KNOWLEDGE GAINED

Identifiable impact craters on Venus are rare—slightly less than one thousand, or approximately 1 crater per million square kilometers—and large craters and basins are uncommon. Impact craters on Venus are randomly distributed and range in size from 1.5 to 270 kilometers in diameter. Venusian impact craters are unusual in that, almost without exception, they appear to be fresh, characterized by sharp rims and well-preserved ejecta deposits. This morphology suggests that the craters have not



A stereo image of Venus's Geopert-Meyer Crater from Magellan. (NASA/JPL) been subjected to significant erosional, volcanic, or tectonic activity. Only about 40 percent of Venusian craters appear slightly modified, 5 percent appear embayed by volcanic deposits, and 35 percent appear modified by tectonic activity. Venus's dense atmosphere filters out small meteors, so there is a lack of small impact craters to chip away at larger craters. This situation favors the preservation of existing large craters. Furthermore, while there is currently no hydrogeologic cycle on Venus, there is evidence to suggest that there may once have been liquid oceans of some kind on the surface.

The pristine appearance of craters on Venus makes it appear the surface is both geologically young and of a relatively uniform age. This observation has significant implications for the geologic history of Venus. While resurfacing processes have most likely removed Venusian craters older that 450 million years, the morphology of existing craters is not what is expected for a steady balance between crater formation and crater loss caused by tectonic, volcanic, or erosional process. The unique observation is that Venusian craters of all ages look "fresh," suggesting that most of the present surface characteristics of Venus date from the end of a global resurfacing event that ceased about 450 million years ago.

Because of their lack of weathering, Venusian craters provide a unique opportunity for scientists to study the effects atmospheric variabilities and gravity have in forming planetary surfaces. While the total number of impact craters on Venus are not comparable to those on Mars, Mercury, and the rocky satellites, they do fall into morphological and age classifications similar to those of impact craters on Earth. Crater density and morphology suggest that cratering records of Venus and Earth are similar. Because the cratering data from these two planets are complementary, they provide interpretive guidelines for researching the roles that volcanism, tectonics, and erosional processes play in planetary resurfacing.

CONTEXT

The high temperatue and dense atmosphere of Venus slow incoming projectiles, destroying the smaller, high-velocity objects. This shielding effect influences the size of Venusian craters. Smaller craters appear to be absent on Venus because only large impactors can penetrate the Venusian atmosphere to reach the surface. It is estimated that as many as 98 percent of the craters between 1.5 and 35 kilometers in diameter that could have formed on Venus did not as a result of its dense atmosphere. Venus's high temperature and dense atmosphere also impeded the emplacement of ejecta during cratering by limiting flight distance and decelerating fragments, resulting in lobate ejecta blankets that are sharply defined and make up coarse blocks. Because of Venus's high surface temperature, rocks tend to be softer, less solid, and somewhat viscous. During an impact, these viscous rocks produce large amounts of impact melt, which works to fill the craters and make them topographically low. It is also suggested that large Venusian impacts could trigger regional tectonic or volcanic events by transferring their tremendous heat and shock energies into the planet's thin crust.

Randall L. Milstein

FURTHER READING

Esposito, L. W., E. R. Stofan, and T. E. Cravens, eds. *Exploring Venus as a Terrestrial Planet: Geophysical Monograph 176*. Washington, D.C.: American Geophysical Union, 2007. Addresses the open questions regarding Venus's geology, atmosphere, surface evolution, and future exploration. Includes results from the Venus Express mission.

Lopes, R. M., and T. K. P. Gregg. *Volcanic Worlds: Exploring the Solar System's Volcanoes*. New York: Springer, 2004. A general review of volcanic activity throughout the solar system. Comparisons are made between volcanic activity on Earth and other planets, showing how data from one planet can aid in the understanding of physical processes on another.

Spudis, P. D. *The Geology of Multi-ring Impact Basins: The Moon and Other Planets*. New York: Cambridge University Press, 1993. Although this is a technical book, the chapter on Venus is well illustrated and easy to read, with good comparisons to similar Earth structures.

Trefil, James S. *Other Worlds: Images of the Cosmos from Earth and Space*. Washington, D.C.: National Geographic Society, 1999. A richly illustrated book with exploration mission images of Venus and computer-generated three-dimensional perspectives of Venus's surface.

Uchupi, E., and K. Emery. *Morphology of the Rocky Members of the Solar System*. New York: Springer, 1993. Focuses on the morphology of planets and their satellites and the reasons for the differences and similarities between them. The book's theme is that the solar system should be approached as a single entity, not as a group of individual planets.

See also: Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Earth-Sun Relations; Extraterrestrial Life in the Solar System; Greenhouse Effect; Mars's Atmosphere; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Rotation; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets; Venus's Atmosphere; Venus's Surface Experiments; Venus's Surface Features; Venus's Volcanoes.

Venus's Surface Experiments

Categories: Planets and Planetology; Venus

An understanding of the geology of the other planets in the solar system is important for understanding the geologic past and future of Earth. Venus holds many clues to this understanding, including its surface geology, which appears to be mostly igneous and basaltic in nature.

OVERVIEW

The planet Venus is considered one of the terrestrial or Earth-like planets because of its position in the solar system, its planetary diameter, its geology, and other characteristics. Despite being called Earth's twin because of those similarities, Venus is actually very different. The study of the planet Venus has been at best difficult because of its heavy cloud cover.

The best information on the Venusian surface and its soils came early from the Soviets, who focused on exploring the planet, successfully landing six spacecraft on the surface. Even though these craft operated for only limited amounts of time because of the planet's extreme temperatures and the pressure of its atmosphere, the data provided from them have given astronomers and geologists important clues to the soils on Venus. Much of the information obtained by the Soviets can be compared with that known about Earth, and to the data obtained from firsthand examination of the lunar rocks and soils. For example, photographs can be useful in examining the appearances of the soil, rocks, and their distribution. Images taken by Venusian landers can be compared with photographs of similar materials found on Earth and the Moon.

The first of the Soviet landers to provide clues to the Venusian soils was Venera 8, which made the first soft landing on Venus, on July 22, 1972, in a region generally thought to be like the rolling plains of Earth. The probe analyzed the surface and soils directly underneath it with a gamma-ray spectrometer designed to determine the chemical composition of surface material. Results showed that the soils under the Venera 8 were igneous, or volcanic, in origin. The layer was found to be approximately 4 percent potassium, approximately 200 parts per million uranium, and approximately 650 parts per million thorium. The layer was also determined to have a density of approximately 1.5 grams per cubic centimeter (in comparison, water has a density of 1 gram per cubic centimeter at a temperature of 277 kelvins). From these data, astronomers and geologists were able to ascertain not only that the soils under Venera 8 were igneous but also that they were probably similar to the granites or basalts found on Earth.

In 1975, Veneras 9 and 10 provided an even better look at the Venusian soils. Each lander transmitted an image that showed the soil and rocks surrounding it. These photographs re-

Venus Compared with Earth

Parameter	*Venus*	*Earth*
Mass (10^{24} kg)	4.8685	5.9742
Volume (10^{10} km^3)	92.843	108.321
Equatorial radius (km)	6,051.8	6,378.1
Ellipticity (oblateness)	0.000	0.00335
Mean density (kg/m^3)	5,243	5,515
Surface gravity (m/s^2)	8.87	9.80
Surface temperature (Celsius)	+450 to +480	–88 to +48
Satellites	0	1
Mean distance from Sun millions of km (miles)	108 (67)	150 (93)
Rotational period (hrs)*	–5,832.5	23.93
Orbital period	224.7 days	365.25 days

*The minus sign signifies a retrograde rotational period.
Source: National Space Science Data Center, NASA/Goddard Space Flight Center.

vealed rocks that were on the average 20 centimeters wide, about 50 to 60 centimeters long, and slablike in appearance. A few of the rocks showed evidence of volcanic origin. Many of the rocks had jagged edges, which demonstrates little erosion, although some did show signs of weathering. This relative lack of erosion surprised many astronomers and geologists. They had believed that, because of the planet's extremes in temperature, atmospheric pressure, wind velocity, and chemical composition, the photographs would show well-eroded landscapes. Astronomer Carl Sagan, among others, hypothesized that low wind velocities at the surface levels of Venus produce little effect on the rock. Apparently, the Venusian surface temperature stays fairly constant and thus does not create much wind. Chemical analysis of the rocks again showed the elements potassium, uranium, and thorium. Nevertheless, the sites differed in the type of rock material. At one Venera lander site, the rocks were basaltic in appearance, similar to those lining Earth's oceans. At the other site, the rocks were more like granite, similar to that found in Earth's mountains. The rocks appear to be relatively young in age. This would indicate that the planet has been geologically active in the geologically recent past. The Venusian soil in the areas observed photographically appeared to be loose, coarse-grained dirt. It was also evident from the photographs that Venus (or at least parts of it) is a dry and dusty planet. Radar images from other Venera missions, as well as Pioneer Venus and the Magellan spacecraft, verified this for the rest of the planet.

The Soviets continued their studies of the Venusian surface with two additional spacecraft, Veneras 13 and 14. These two spacecraft performed similar examinations, but in a much more complex manner. Rather than single images, near-panoramic views of the landing sites were produced. Photographs showed rocks somewhat similar to those found at the Venera 9 and 10 landing sites. Rocks also showed evidence, however, of what appears to be thin layering, ripple marks, and fracturing, especially around Venera 14. Some rocks showed evidence of erosion. On Earth, rocks that show layering—such as sandstone and limestone—are usually sedimentary. Based on the photographs and measurements made by the spacecraft, several Soviet scientists suggested that the Venusian rocks might be sedimentary, but that has not been confirmed.

The possible cause or causes of the erosion remain unknown. In the absence of water, several possibilities have been suggested. These include chemical weathering or erosion caused by nearby volcanism and its resulting ash, dust, and lava. Chemical weathering seems the most likely explanation. Venus's thick atmosphere has cloud layers laced with sulfuric acid, which rains down as a caustic, corrosive agent on the surface.

Both spacecraft collected a cubic centimeter of Venusian soil for analysis. The probes utilized an X-ray source to stimulate emissions from the collected soil samples. This chemical analysis revealed that the samples were similar to basalt in composition, although the basalts differed at the two sites. Near the Venera 13 landing site, the type of basalt found is referred to as leucitic high-potassium basalt, while near the Venera 14 landing site, a tholeiitic basalt, similar to that found on the ocean floors on Earth, was found. The soil itself appeared fine-grained, and the photographs revealed many small rocks. It has been speculated that this also indicates that weathering processes of some type are at work, breaking down larger rocks into smaller ones, eventually reducing them to soil.

Another pair of Soviet probes, the Vega 1 and 2 spacecraft, landed on Venus in June, 1985. Vega 2 results revealed a Venusian soil and surface that are again similar to basalt. Nevertheless, the new data also revealed a surface rich in the element sulfur, which is usually associated with volcanism. This presence has provided another clue to the surface and geology of Venus.

Venera 8 landed about 5,000 kilometers east of an area referred to as the Phoebe region. Veneras 9, 10, 13, and 14 all landed between 900 and 3,000 kilometers east of the raised areas known as Beta Regio and Phoebe Regio. Even though the craft landed on and took samples from an area that could be of the same or similar geologic makeup, they have given astronomers and geologists a good idea of the planet's surface composition.

The probes produced mostly photographic data, although some chemical analysis was conducted on site. Thus, any discussion of soil samples is based on the evidence reported by these spacecraft, since no samples have ever been returned to Earth for detailed study. Spacecraft data have enabled astronomers and geologists to begin to understand not only the surface of the planet Venus and its chemical makeup but also the planet's evolutionary path.

KNOWLEDGE GAINED

It appears that Venus may still be an active planet geologically, which scientists inferred from the discovery of high concentrations of sulfur in Venus's atmosphere. Thus, its soils, for the most part, must be considered with that fact in mind.

Analysis of Venusian rocks around the landers provided scientists with interesting but sometimes confusing data. For example, the fact that most of the rocks appear to lack signs of erosion at first seemed puzzling. An understanding of the weather patterns on Venus and the planet's atmospheric chemistry, however, has led to the development of theories relating the small-scale erosion to a low wind velocity at the surface because of its virtually uniform temperature. The rocks themselves appeared to be mostly igneous in nature. Most igneous samples appeared to be similar to basalt, much like those rocks and materials that line Earth's ocean floors. Some of the rocks resembled granite, like those that form Earth's mountains. However, despite apparent volcanic origins for Venus's crust, some specimens appeared to be sedimentary. This led to further questions which remain unanswered. Although the sedimentation process on Earth is usually accomplished by water, present-day Venus has no water, nor is there any evidence of water in its near past. The origins of this phenomenon remain unknown.

Analyzed samples varied slightly from site to site, as was expected by geologists, since samples on Earth also differ. In fact, variation of Earth samples is greater than that of the limited Venusian ones. Nevertheless, potassium—a key element in igneous and especially basaltic materials—was detected, as were uranium and thorium. Geologically, the rocks are relatively young, presenting additional evidence that Venus is a planet that may be experiencing continuous changes. Fine-grained soils were found at some sites, while coarser soils appeared at others. At one site, at least, smaller rocks led scientists to theorize that erosion does occur on the planet, thus producing soil.

CONTEXT

When the planets of the solar system are categorized, one usually finds two major groupings: the terrestrial or Earth-like planets; and the Jovian or Jupiter-like planets (sometimes also called the "gas giants"). Venus, because of its relative size, atmosphere, position within the solar system, and surface, is naturally among the set of terrestrial planets: Mercury, Venus, Earth, and Mars.

An understanding of the nature of the terrestrial planets, their atmospheres, planetary geologies, and soils can give astronomers and geologists clues to the pasts not only of these worlds but also of our own—revealing how these planets were formed, what geological changes they have undergone, and how they might be related. Venus holds many clues to the formation of the solar system. Unfortunately, observations of the Venusian surface are nearly impossible because of the dense atmosphere that surrounds the planet. Orbiting spacecraft provide information regarding the general geologic contours on the surface—the planet's mountains, valleys, craters, and other surface features—but hard evidence of the nature of the surface, particularly its soil, can come only from the surface of the planet. Prior to the landing of Soviet probes, no information about the Venusian surface existed.

Materials sampled and photographed in the vicinities of the Soviet landers proved to be mostly igneous in nature. Additional on-site chemical analysis showed these materials—both rocks and soils—to be similar to granite or basalt. Basalt-type materials are not unique to the second planet from the Sun. These materials have been found on Mars in the vicinities of the American Viking landers, in samples brought back from the Moon by the American Apollo crews, and by uncrewed Soviet spacecraft. As Soviet spacecraft became more sophisticated and knowledge of the harsh Venusian environment grew, landers were able to provide data on the surface of Venus, among other things. Additional information may provide scientists with clues to the past of the terrestrial planets, part of which is hidden in the Venusian surface and soil. Perhaps more important, Venusian soil information may provide clues to Earth's future, particularly regarding our planet's fragile environment.

Comparative planetology is essential for achieving a more complete understanding of our own planet. As Venus, Earth, and Mars started out relatively similar in the early solar system, and all three are in the habitable zone, why is it then that Venus is devoid of water and hot with a thick atmosphere of carbon dioxide, Earth is capable of supporting life, and Mars has no liquid surface water and is cold, with a thin atmosphere of carbon dioxide? Only when that question is answered will scientists have a clear idea of Earth's complex planetary environment.

Work performed along the way to that understanding has included spacecraft dispatched to the veiled planet Venus by both the Soviet Union and the United States. American spacecraft have only flown by or Venus or studied it from orbit. The National Aeronautics and Space Administration (NASA) has had more interest in exploring Mars than Venus. However, because the Soviets have had only bad luck when it comes to Martian exploration, they have emphasized the study of Venus with flyby craft, landers, orbiters, and even balloons temporarily floating within its hellish atmosphere. At the dawn of the twenty-first century, however, neither NASA nor the Russians had any plans to return to Venus for at least two decades. In the meantime, the European Space Agency's Venus Express began orbiting Venus on April 11, 2006, conducting mapping operations and other scientific investigations of Venus's surface and atmosphere.

Mike D. Reynolds

FURTHER READING

Cattermole, Peter John. *Venus: The Geological Story*. Baltimore: Johns Hopkins University Press, 1996. A comprehensive presentation of the latest findings on Venus, based on Magellan data.

Corliss, William R., ed. *The Moon and the Planets*. Glen Arm, Md.: Sourcebook Project, 1985. A discussion of many solar-system phenomena that cannot be easily explained by prevailing scientific theories. Each anomaly is defined, substantiating data are presented, and the challenge the anomaly presents to astronomers is explained. Examples and references are also listed.

Esposito, Larry W., Ellen R. Stofan, and Thomas E. Cravens, eds. *Exploring Venus as a Terrestrial Planet*. New York: American Geophysical Union, 2007. A collection of articles covering all major areas of planetary research on Venus. Technical.

Fimmel, Richard O., Lawrence Colin, and Eric Burgess. *Pioneering Venus: A Planet Unveiled*. Washington, D.C.: National Aeronautics and Space Administration, 1995. A profusely illustrated scientific and technical publication from NASA that includes the Pioneer Venus data as well as a good deal of information from the Russian spacecraft dispatched to investigate Venus. Illustrations, bibliographic references, index.

Frazier, Kendrick. *Solar Systems*. Rev. ed. Alexandria, Va.: Time-Life Books, 1985. This text contains outstanding color photographs, diagrams, and coverage of the planets of the solar system.

Grinspoon, David Harry. *Venus Revealed: A New Look Below the Clouds of Our Mysterious Twin Planet*. New York: Basic Books, 1998. A thorough examination of the geology of Venus. Incorporates Magellan mapping and other data. Explains the Venusian greenhouse effect. Includes speculation about Venus's past. A must for the planetary science enthusiast who wants an integrated approach to science and history.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. An updated version of a classic text that covers all areas of planetary science. The chapter on Venus covers all fundamental knowledge about the planet and spacecraft exploration of it.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. The authors provide a detailed description of each of the planets and other bodies of our solar system. Some coverage of general astronomy and chemistry is included in introductory chapters. College level.

Snow, Theodore P. *The Dynamic Universe*. Rev. ed. St. Paul, Minn.: West, 1991. A general introductory text on astronomy. Covers historical astronomy, equipment used in astronomy, the solar system, stellar astronomy, galactic astronomy, cosmology, and life in the universe. The book features special inserts, guest editorials, and a list of additional readings at the end of each chapter. College level.

See also: Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Earth-Sun Relations; Extraterrestrial Life in the Solar System; Greenhouse Effect; Mars's Atmosphere; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Rotation; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets; Venus's Atmosphere; Venus's Craters; Venus's Surface Features; Venus's Volcanoes.

Venus's Surface Features

Categories: Planets and Planetology; Venus

Enormous strides have been made in understanding the nature of the surfaces of the solid bodies in the inner solar system and the processes that have shaped them. Venus, however, has been a particularly difficult planet to study. The picture that is emerging suggests that Venus may be the only member of the four terrestrial planets, besides Earth, that remains geologically active.

Overview

Of all the terrestrial planets, Venus is the most similar to Earth in size and geologic composition. At 12,258 kilometers in diameter, it is only slightly smaller (by 511 kilometers) than Earth, and its density is within 2 percent of being identical. It lacks the polar flattening, equatorial bulge, and planetary magnetic field that Earth exhibits.

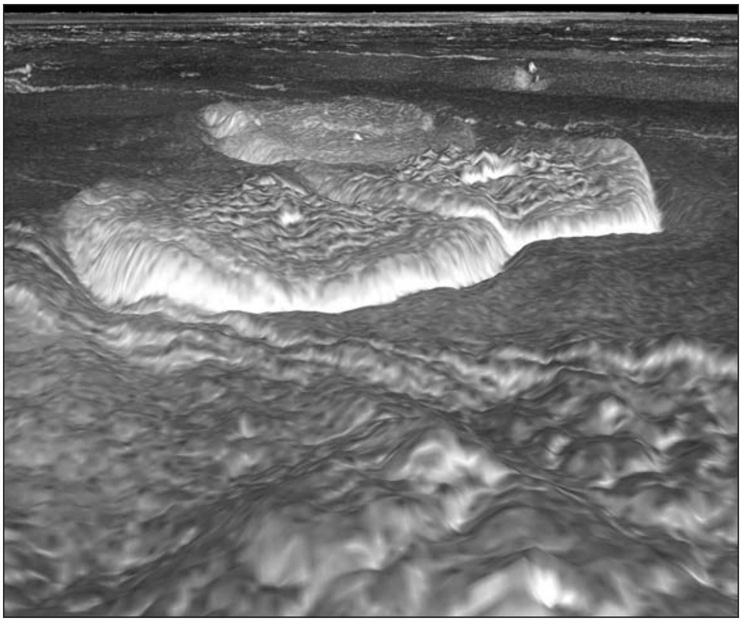
Geologic study of Venus is exceedingly difficult because of the fact that the planet is perpetually shrouded from view by thick clouds. Its surface has never been photographed from Earth-based telescopes or from spacecraft in orbit above the planet. However, in 1990 the Magellan spacecraft began generating high-resolution radar images of Venus's surface and thereby began producing detailed maps with a resolution of approximately 100 meters. A few panoramic photographs taken by several Soviet spacecraft have revealed the barren, rocky character of Venus's surface in the proximity of their landing sites. However, a very high surface temperature averaging 750 kelvins has limited the operating life spans of spacecraft that have landed on Venus to a maximum of about two hours. Still, Venus has been the subject of very persistent research by space scientists and has yielded enough data about its topography and composition to permit informed speculation about the processes responsible for creating its surface. Scientists' knowledge of the planet's geologic features rests primarily on techniques involving radar imaging, while preliminary impressions of the chemical and structural nature of the surface have been provided through experiments conducted by Soviet spacecraft at several different landing sites.

The Venusian surface is generally smoother than that of any other terrestrial planet. Sixty percent of it lies within 500 meters of Venus's mean radius of 6,051 kilometers. Because Venus has no equivalent to sea level, the mean radius is used as the baseline elevation for topographic measurements. Despite this prevailing uniformity, Venus does have some high mountains and deep valleys. The total range between highest and lowest points on the planet is nearly 14 kilometers, a value that is similar to Earth's.

Planetologists divide Venusian topography into several distinct types of terrain. Rolling plains dominate the globe and form an irregular, planet-girdling area covering more than 70 percent of the surface. About 16 percent of the surface lies below the level of the rolling plains. The remainder is divided among upland plains (0.5-2.0 kilometers above mean radius) and several types of true highlands. Among the latter are the regios, also called domed uplands. These are large, roughly circular areas that rise gently toward their centers, where they achieve heights of between 3 and 5 kilometers above the mean radius. They are thought to be situated over interior "hot spots," which have caused the surface to bubble outward on a gigantic scale. Huge shield volcanoes sit atop many of the domed uplands, a fact that adds credibility to the theory that these landforms are similar to volcanic domes on Earth. Alpha Regio and Beta Regio, the first two surface features identified by Earth-based radar studies, are examples. Surfaces of the domed uplands are generally smooth, like those of the rolling plains, and appear to be the same age. Unlike plains, however, they seem to be crisscrossed by fault lines indicating crustal stresses. Two continent-sized highland areas lie within the mapped region of the surface, but together they account for only 8 percent of the planet's surface. Ishtar Terra and Aphrodite Terra, about the size of Australia and Africa respectively, exhibit a rich variety of landscapes, including two types of mountainous terrain as well as areas of flat and complex plains.

The most common highland topography is one of ridges and valleys that intersect in chevron-shaped or chaotic patterns. This terrain is called tessera terrain, from the Greek word for "mosaic tile," and resembles the deformation patterns that occur on the top of a moving glacier; there are no glaciers on Venus, however. In contrast to the tesserae are more dramatic but less common mountain systems that thrust their peaks 4 to 12 kilometers above the mean radius. The Maxwell Montes region of Ishtar Terra, which includes the highest known point, 11,800 meters above mean radius, is an example

In 1975, the Venera 10 lander survived for sixty-five minutes on the Venusian surface and took this picture of a volcanic-looking surface and part of the spacecraft before it lost function. (NASA)



Part of Alpha Regio, a mountainous region with seven domed hills, likely built by outflows of lava. (NASA/JPL)

of the latter. It consists of a series of parallel ridges and valleys 15 to 20 kilometers apart. The Maxwell system appears much like the ridge and valley province of the Appalachian Mountains, although it is substantially higher. Its features, and those of at least three other mountain chains on Ishtar Terra, closely resemble those produced when plates of the Earth's crust are thrust together by tectonic forces.

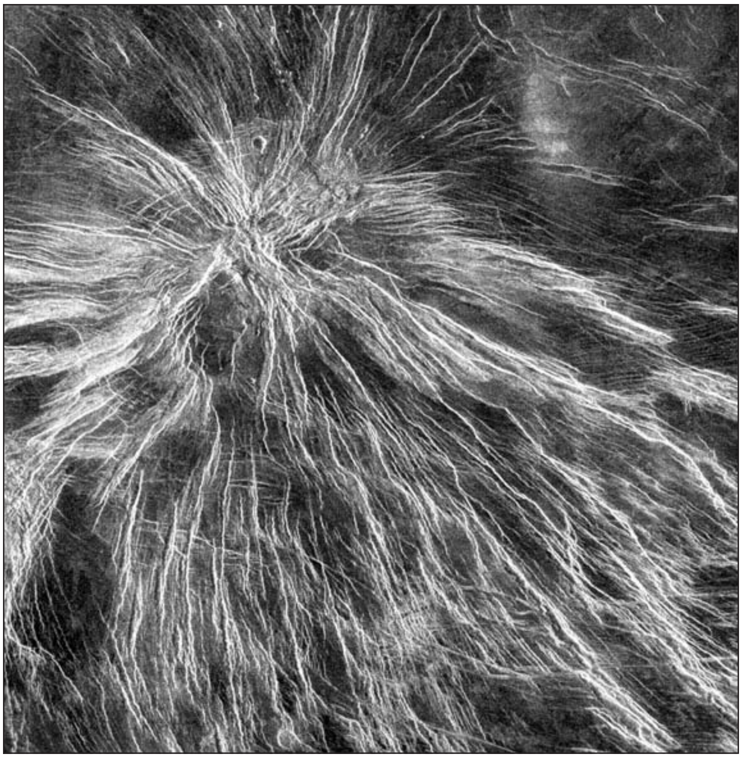
Contrasting with the mountain ranges are great rifts that cleave the surface to depths of up to 2 kilometers. One huge rift system stretches from east to west for more than 20,000 kilometers and can be traced as a series of chasms along the entire southern edge of Aphrodite Terra. From there, it continues across the rolling plains to link with Beta Regio. Another rift splits Beta Regio and continues to Phoebe Regio. This complex system consists of many related but distinct chasms, the largest of which is 3,500 kilometers long and 100 kilometers wide, with its deepest point lying 2.1 kilometers below mean radius.

Among the most interesting surface features thus far discovered are Venus's large volcanoes. Two excellent examples are Thea and Rhea Mons, which, along with several other volcanoes, rise from the Beta Regio dome. They appear to be situated on a fault that forms one edge of the great rift. The mass of material that has issued from them is greater than the total output of the volcanic mountains that have formed the Hawaiian Island chain. Both are shield volcanoes like Olympus Mons on Mars, formed by chronic, nonexplosive eruptions of lava that flow long distances before solidifying. Radar images show what may be geologically recent lava flows from both Thea and Rhea, and there is intriguing but very controversial evidence that these volcanoes may have been in eruption in the 1950's and again in the 1970's. Since that time no direct evidence of volcanism has been found by even the long-lived Magellan orbiter.

Impact craters, so characteristic of the surfaces of the Moon and Mercury and even fairly common on Mars, are apparently not nearly so plentiful on Venus. More than one hundred have been observed in radar images, ranging in diameter up to a maximum of 144 kilometers. However, the number of craters discovered is considered low, and the largest crater is only modest in size. These facts are considered to be important evidence that the present Venusian surface is not an ancient one. An old surface should bear numerous scars of past encounters with meteors, comets, and asteroids, as is the case with the Moon and Mercury.

Another class of circular geologic feature bearing superficial similarity to impact craters is apparently an unrelated phenomenon. This group comprises the so-called coronae, of which at least eighty have been found. They average 500-800 kilometers in diameter, but their depth is only 200-700 meters, a fact that is inconsistent with an impact origin. Many researchers interpret the coronae to be collapsed bubbles in the crust, caused by localized heating from hot spots in the mantle beneath.

Many Soviet spacecraft failed, but four successfully took panoramic photographs of their



Graben—long, linear depressions usually found between parallel faults—are visible in Venus's Themis Regio. These particular graben form a nova, or a series of graben radiating from a central area; about fifty such novae have been found on the Venusian surface. (NASA)

landing sites. These pictures are remarkably similar in showing barren landscapes dominated by flat-topped rocks with relatively little loose, fine-grained material that might be described as soil. Closer inspection of images of the rocks shows that they seem to be partially exposed outcroppings of a horizontally layered rock mass that exhibits a marked tendency to break into platelike slabs. Additional measurements indicate that the rocks are of low density (1.5 grams per cubic centimeter) and high porosity. They have a bearing strength of only a few kilograms per square centimeter, meaning that they can be broken rather easily. These findings are regarded as surprising, for they are characteristic of sedimentary rocks on Earth. Under close inspection, the panoramic photographs seem to support this conclusion, showing what appear to be striations, indicating ripple marks and crossbedding, two common features of sedimentary deposits. Even the electrically nonconductive properties of the rocks, as revealed by their radar reflectivity, agrees well with the behavior of sedimentary rock. If the surface rocks are indeed of sedimentary origin, they presumably formed from deposits of windblown sand.

The velocity of surface winds is low by terrestrial standards, not exceeding 1.3 meters per second. However, under Venus's dense atmosphere, which is ninety times heavier than Earth's, this velocity is more than sufficient to move fine-grained materials and raise dust. At a number of sites on the rolling plains, researchers have detected depressions that seem to be filled with volcanic ash, which may become lithified over time by yet unknown processes.

The mass of the planet, the density of its surface materials, and the relative abundance of certain elements present in its rocks all point to the likelihood that Venus, like Earth, experienced a planetwide "meltdown" early in its history. The result was differentiation, a process in which the lighter elements migrated to the surface and the heavier elements settled toward the center. Escape of the residual heat from that meltdown is presumed to have been the major architect of the surface features observed on Venus, just as it has been on Earth. Whether the interior remains molten has not been determined. Venus unquestionably lacks a planetary magnetic field; because such a field is thought to be generated by planetary rotation around a molten iron core, this lack suggests that the interior of Venus has cooled and solidified. However, the evidence that volcanism has occurred on the surface in recent geologic time contradicts this view. It may be that Venus, which rotates 243 times more slowly than does Earth, simply does not spin fast enough to create the dynamo effect that gives rise to a magnetic field.

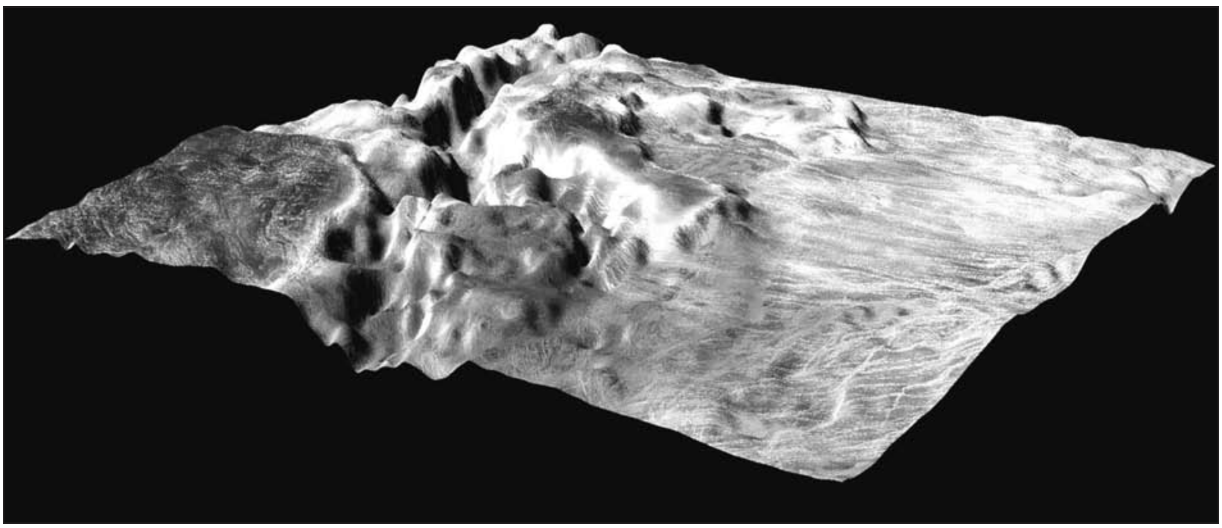
METHODS OF STUDY

Study of Venus by optical telescope has not been productive for elucidating the nature of the planet's surface. The first successful attempts to penetrate Venusian clouds were made in 1961. Teams of American, British, and Soviet scientists were able to use large radio telescope antennas to beam radar waves at the planet and receive faint return echoes. Earth-based radar studies of Venus have continued but are seriously limited by the fact that good results can be achieved only when the planet passes near the Earth, at which time Venus always presents the same "face."

To gain a global picture of the Venusian terrain, the National Aeronautics and Space Administration's (NASA's) Pioneer Venus orbiter and the Soviet Venera 15 and 16 spacecraft carried radar-imaging instruments into orbit around Venus. In principle, the American and Soviet spacecraft operated similarly, combining synthetic aperture radar (SAR) imaging with radar altimeter measurements. The SAR images from Pioneer Venus cover 70 percent of the surface but are only about as detailed as those shown by a desktop physical relief globe. Veneras 15 and 16 reached Venus in late 1983. Equipped with larger radar antennas, they were capable of resolving surface details as small as one to two kilometers in size and could measure elevations to within 50 meters. Venera radar images look remarkably like high-altitude black-and-white photographs. Unfortunately, this imaging was obtained only for the northern quarter of the planet (from 90° to 30° north latitude).

Beginning in late 1970, a series of soft landings on Venus were made by Soviet Venera craft equipped to conduct experiments to detect the presence of certain rock-forming minerals. One approach used a gamma-ray spectrometer to detect gamma radiation emitted by radioactive uranium, thorium, and potassium. Venera landers of the 1980's employed an automated drill that bored into the surface to obtain samples not contaminated by chemicals from the atmosphere or from the lander itself. The sample material was transferred to an automated laboratory inside the craft, where it was subjected to X-ray fluorescence. Results showed that the rocks at most of the landing sites appear to resemble basalt, an igneous rock enriched with iron and magnesium. The exact composition, however, varied from site to site and, while not identical to that of Earth basalts, was chemically closer to Earth rocks than to Moon rocks.

At two locations, the experiments detected minerals more characteristic of granite, another common igneous rock. These findings are not seen to be in conflict with evidence that Ve-



A computer-generated image from Magellan shows a portion of the Aphrodite Terra called Ovda Regio, where plains meet highlands. (NASA/JPL/USGS) nusian surface rocks may be sedimentary in nature, for the experiments detected the presence and ratio of identifying minerals but not the type of matrix that contained them.

There is no water on the surface of Venus at present, but two discoveries suggest that such has not always been the case. If water was ever plentiful, it must have boiled away, so that the water molecules were dissociated into oxygen and hydrogen. Investigators have sought evidence for the "missing" oxygen and hydrogen, and some believe that they may have found both. Deuterium, a hydrogen isotope, has been detected to be one hundred times more abundant in the Venusian atmosphere than it is in Earth's. Meanwhile, an experiment has shown that oxidized terrestrial basalts, when heated to the Venusian surface temperature, appear identical in visible and micrometer wavelength imagery to the surface rocks of Venus. The likeliest source for the deuterium and the oxygen to oxidize the basalt is dissociated water molecules.

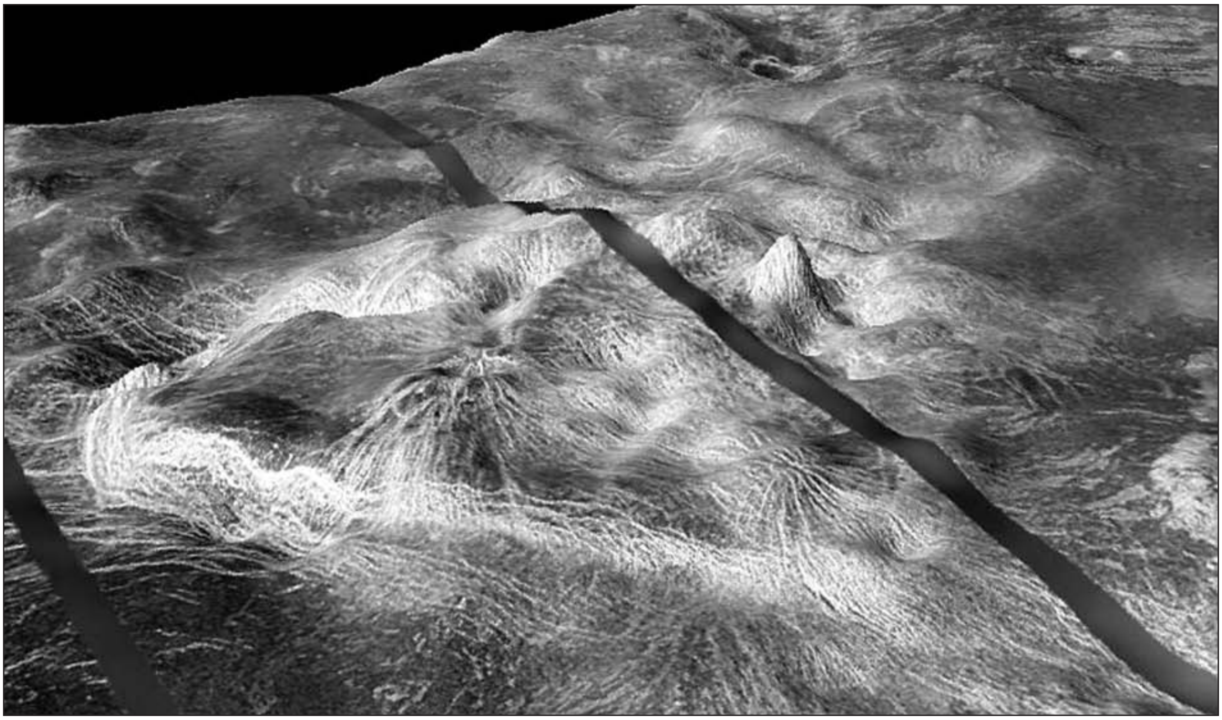
An intriguing possibility exists that Venus may harbor active volcanoes—perhaps the largest in the solar system (the large Martian volcanoes are definitely extinct). Spacecraft and Earth-based observations have detected large amounts of sulfur dioxide, a common volcanic effluent, in the Venusian atmosphere. Moreover, sulfur dioxide content increased dramatically in the 1950's and again in the 1970's.

The first color pictures taken on the surface by the Venera 13 lander seemed to show that the landscape had an orange or amber tint, which proved to be an effect of sunlight filtered through the heavy overcast. Computer processing of the photographs has since shown that, in normal white light, the rocks are a uniform, colorless gray.

Previous spacecraft had performed preliminary radar investigations of Venus's surface, identifying the major types of features on that surface and identifying prominent examples of each. However, high-resolution maps of the entire surface were lacking. The goal of NASA's Magellan spacecraft was to use a synthetic aperture radar in prolonged orbit about Venus to produce a global map at a resolution even in excess of the best contemporary maps of Earth's surface. Detailed geological interpretations and altimetry data were obtained in the process.

Magellan launched aboard the space shuttle Atlantis on May 4, 1989, and was the primary payload of the STS-30 mission. The spacecraft was deployed from the shuttle's cargo bay, and was dispatched on an trajectory that concluded with orbital insertion about Venus on August 10, 1990. Magellan entered a highly elliptical orbit, often ranging from as little as 300 to as much as 8,500 kilometers above the surface. With Magellan in a polar orbit, the planet Venus rotated underneath it, thereby allowing the spacecraft to image a different ground track on each low pass. The spacecraft turned toward Earth as it climbed toward its highest orbital point and then transmitted the radar imagery it had collected during its low pass over Venus.

Science activities were slightly varied with each of Magellan's six cycles, with the spacecraft's orbit occasionally being altered for different research requirements. During the first cycle, Magellan concentrated on global radar mapping and imaged 84 percent of Venus. Later cycles filled in gaps and concentrated on specific features of interest. In a lower orbital altitude late in its operational mission, Magellan was able to collect precise gravitational data as Venus slightly altered the spacecraft's orbital parameters. Magellan was used to test aerobraking techniques by having the spacecraft fly through the upper portions of Venus's atmosphere; its large solar panels experienced a retarding torque due to atmospheric drag. How the spacecraft responded to the atmosphere indirectly informed scientists about Venus's atmospheric particle density as a function of altitude. With its primary and extended missions completed, flight controllers decided to send Magellan plunging into the upper atmosphere to remove it from orbit. Maneuvers were conducted to force the spacecraft's orbit to decay due to orbital drag. On October 11, 1994, the final spacecraft maneuver was conducted. Controllers lost contact with Magellan the following day. Then, on October 14, Magellan was destroyed in the atmosphere. Although it could not be verified, many believed pieces of descending debris survived long enough to impact the surface.



Yavine Corona, one of many coronae, circular regions averaging 500-800 kilometers in diameter, possibly collapsed crustal bubbles caused by localized heating from hot spots in the mantle beneath. Yavine Corona contains two novae. (NASA/JPL/USGS)

Perhaps the most important results of Magellan's intense investigation of Venus was determining a total lack of plate tectonics based on the two primary processes observed on Earth. Instead of continental drift and basin floor spreading, Venus's global rift zones and coronae move as a result of upwelling and subsidence of magma in the planet's mantle. That suggested that Venus's surface is indeed quite young geologically speaking, perhaps less than 800 million years old. The enormous data set from Magellan was made available to interested researchers and individuals on compact disc.

Despite the extensive research with Magellan, many questions remained to be investigated. The European Space Agency (ESA) dispatched its Venus Express spacecraft to Venus in order to examine the planet's atmosphere at infrared wavelengths. Venus Express launched on November 9, 2005, and was inserted successfully into Venus orbit on April 11, 2006. Its Visible and Infrared Thermal Imaging Spectrometer (VIRTIS) was used to identify the amount of sulfur dioxide in the atmosphere between 35 and 40 kilometers and to monitor that constituent for changes in concentration over time that would indicate active volcanism. Venus Express's Spectroscope for Investigation of Characteristics of the Atmosphere of Venus (SPICAV) used stellar occultation methods to determine the identity of atoms and molecules in the upper atmosphere at an altitude between 70 and 90 kilometers. SPICAV saw rapid drops in the amount of sulfur dioxide in the upper atmosphere, strongly indicating that Venus has active volcanoes. VIRTIS was then used to identify hot spots on the surface.

CONTEXT

In order to understand a system as complex as a terrestrial planet, it is necessary to have more than one example of how such planets function and evolve. Hence, the study of geologic processes on another world, far from being simply an esoteric and impractical inquiry, holds promise for improving human understanding of the forces that have acted on Earth and still continue to shape its surface. For this reason, the goal of Venusian geological studies is to discover the relationships and sequences of events that have resulted in the landforms that can be imaged. "Looking at Venus is like running the experiment that produced the Earth a second time," according to Robert Kunzig, senior editor of *Discover* magazine. Indeed, scientists appreciate the opportunity to "run the experiment" again under slightly different conditions in order to see whether their set of explanatory theories can accommodate any observed deviations in the results. Venus presents a marvelous opportunity to test the plate tectonics theory on a planet that has many fundamental similarities to Earth but also exhibits numerous significant differences.

It is generally believed that Earth's loss of internal heat is primarily a result of convection and occurs mainly along the 75,000-kilometer-long mid-ocean ridge. Seafloor spreading that results is responsible for producing a large expanse of young and renewable crust. Even the older continental masses are invigorated by the tectonic activity that is driven by this convective heat loss. Venus seems to have experienced similar horizontal crustal movements in the past and may still be experiencing them. The driving force, however, might be quite different. Most authorities interpret the present surface as having been formed through the release of heat at localized hot spots.

Another fundamental question is whether rocks making up Venus's vast, rolling plains differ significantly in composition from those of the higher terrain. This issue is of interest because it bears on where and how crustal materials originated. It is not inconceivable that Venus could reveal hitherto unknown relationships that may have acted on Earth when the original continental rocks were first solidifying some 3.8 billion years ago.

The mystery of whether Venus once had oceans is of particular interest for two reasons. First, most scientists now accept as true evidence that Earth's own atmosphere is beginning to warm as a result of increases in carbon dioxide content. There is growing concern that atmospheric pollution may trigger a runaway "greenhouse effect," similar to the process that appears to have happened on Venus. If Venus retains any "memory" of conditions before it became so hot, it will only be in the record of the rocks themselves. Second, scientists are still uncertain about how Earth got its abundant water in the first place. If Venus had oceans at a former time, that fact would have significant implications for theories of the origin of planetary water.

A large amount of the Venusian surface was mapped by Pioneer Venus's radar, but its images were good enough only to show gross features of the surface and could not address cause-and-effect relationships. Better imaging has been obtained by the Arecibo radio telescope and the Soviet Venera 15 and 16 spacecraft, but the total area covered was too small to permit generalizations to be drawn. Together, all these data allowed planetary scientists the luxury of asking better questions, which might then be answered by the higher-resolution Magellan spacecraft's synthetic aperture radar imaging system. Magellan produced a spectacular increase in knowledge of Venusian topography. Magellan imagery provided planetary scientists with the information needed to correlate the roles of volcanic activity, tectonic motion, and impact events in the formation and evolution of Venus's surface features. Magellan established that some surface features resulted from tectonics and Venusian volcanoes have been active in recent geologic time, but some key questions are likely to remain unanswered until more complicated surface experiments can be conducted.

Richard S. Knapp

FURTHER READING

Bazilevskiy, Aleksandr T. "The Planet Next Door." *Sky and Telescope* 77 (April, 1989): 360-368. The author, a senior member of the Soviet Venera science team, provides a comprehensive overview of the surface of Venus in clear and nontechnical terms. The article is particularly valuable for its discussion of Venus's medium- and small-scale surface features and for its fair-minded discussion of issues about which there is significant debate or uncertainty.

Bredeson, Carmen. *NASA Planetary Spacecraft: Galileo, Magellan, Pathfinder, and Voyager*. New York: Enslow, 2000. This book, part of Enslow's Countdown to Space series, provides an overview of NASA planetary exploration during the last two decades of the twentieth century. Suitable for all audiences.

Burgess, Eric. *Venus: An Errant Twin*. New York: Columbia University Press, 1985. Probably the general reader's most complete single source of information about Venus and how present knowledge has been obtained. Chapters on the Veneras, Pioneer Venus, and the relationship of Venus's geological history to that of Earth and Mars round out the detailed discussion of the surface and atmosphere. Nearly one hundred well-chosen photographs and diagrams illustrate the text.

Cattermole, Peter John. *Venus: The Geological Story*. Baltimore: Johns Hopkins University Press, 1996. Provides a comprehensive presentation of the latest understanding of Venus, based on Magellan data.

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001. A challenging and thorough text for students of planetary geology, this volume offers an excellent reference for the most serious reader with a strong science background. Provides an in-depth contemporary explanation of solar-system formation and evolution.

Faure, Gunter, and Teresa M. Mensing. *Introduction to Planetary Science: The Geological Perspective*. New York: Springer, 2007. Designed for college students majoring in Earth sciences, this textbook provides an application of general principles and subject material to bodies throughout the solar system. Excellent on comparative planetology.

Fimmel, Richard O., Lawrence Colin, and Eric Burgess. *Pioneering Venus: A Planet Unveiled*. Washington, D.C.: National Aeronautics and Space Administration, 1995. A profusely illustrated scientific and technical publication from NASA that includes the Pioneer Venus data as well as a good deal of information from the Russian spacecraft dispatched to investigate Venus.

Grinspoon, David Harry. *Venus Revealed: A New Look Below the Clouds of Our Mysterious Twin Planet*. New York: Basic Books, 1998. Incorporates Magellan mapping and other data in its coverage of Venusian geology; explains the Venusian greenhouse effect; speculates about Venus's past. A must for the planetary science enthusiast who wants an integrated approach to science and history.

Harvey, Brian. *Russian Planetary Exploration: History, Development, Legacy, and Prospects*. New York: Springer, 2007. Early Russian space programs attempted a large number of Moon, Venus, and Mars investigations. Many were successful, many not. These robust programs are often overlooked. This is their story in one illuminating book about the engineering, development, flight operations, and science returns.

Kerr, Richard. "Venusian Geology Coming into Focus." *Science* 224 (May 18, 1984): 702-703. A brief, easily understood summary of the major geologic features revealed by radar mapping. Interpretation of the Venera 15 and 16 results by leading American planetologists forms the basis of the article. *Science* frequently publishes articles dealing with research on Venus. Many are suitable for lay readers and those with a general science background.

Marov, Mikhail Ya, and David Grinspoon. *The Planet Venus*. New Haven, Conn.: Yale University Press, 1998. Marov was Soviet Venera mission chief scientist, Grinspoon a NASA-funded scientist studying Venus. Together they provide a coordinated description of American and Soviet attempts to learn the secrets of Venus, a planet shrouded in mystery. For both general readers and specialists.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003. A full chapter is devoted to the geologic and atmospheric processes on Venus. An additional chapter covers the origin of the solar system, useful to those not familiar with current theories on planetary formation. Each of the other terrestrial bodies also receives a chapter of discussion.

Spanenburg, Ray, and Kit Moser. *A Look at Ve-*

nus. New York: Franklin Watts, 2002. A look beneath the thick clouds of Venus. Written for a younger audience.

Young, Carolynn, ed. *The Magellan Venus Explorers' Guide*. Pasadena, Calif.: Jet Propulsion Laboratory, California Institute of Technology, National Aeronautics and Space Administration, 1990. Prepared as a field and educational guide to the Magellan mission. Published prior to mission launch, this volume contains no spacecraft results, but it does describe the expected research and its value to Venus studies.

See also: Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Earth-Sun Relations; Extraterrestrial Life in the Solar System; Greenhouse Effect; Mars's Atmosphere; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Rotation; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets; Venus's Atmosphere; Venus's Craters; Venus's Surface Experiments; Venus's Volcanoes.

Venus's Volcanoes

Categories: Planets and Planetology; Venus

The planet Venus has at least sixteen hundred major volcanoes and many more minor ones, which is more volcanoes than any other planet in the solar system. Most are shield volcanoes, but Venus also has pancake domes and other volcanic features. About 80 percent of Venus's surface has been shaped by some type of volcanic activity. Venus does not have volcanic chains like those formed on Earth from plate tectonics. Comparing volcanic features on Venus and Earth helps us better understand volcanic processes on both planets.

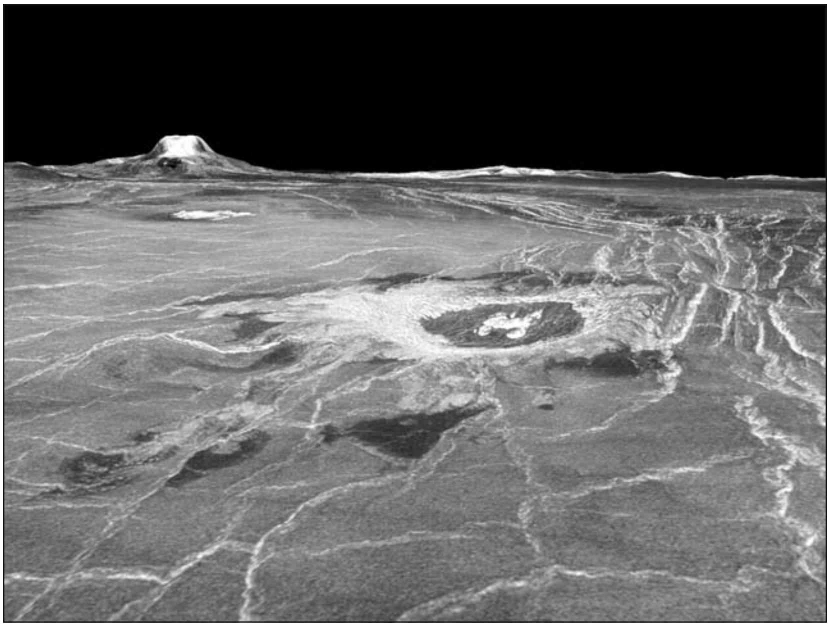
OVERVIEW

Venus is sometimes referred to as Earth's twin sister because its size, mass, and density are similar to those of Earth. Venus's surface conditions, however, definitely make Venus Earth's "evil" twin sister. Owing to a runaway greenhouse effect from the carbon dioxide atmosphere, Venus has a surface temperature hot enough to melt lead. The surface atmospheric pressure is nearly one hundred times what it is on Earth. Thick layers of sulfuric acid clouds veil the surface of Venus. Landers on Mars can last for years, but on Venus they are destroyed by the harsh surface conditions within about an hour. These atmospheric conditions make it impossible to study the surface of Venus using direct optical means or long-term robotic landers. Astronomers must use radar maps rather than optical photographs to study the planet's surface features. Radar maps from both Earth and spacecraft have, however, unveiled the surface of our mysterious twin sister.

Radar maps show that volcanic activity has played a major role in shaping the surface of Venus. Volcanic activity includes not only erupting volcanoes but also lava flows and other activity whereby solid, liquid, or gaseous material escapes from the planet's interior. Volcanic activity is often caused by tectonic activity but can occur independently of tectonic activity. There are more than sixteen hundred large volcanic features on the surface of Venus and possibly as many as hundreds of thousands of smaller volcanic features. In addition, about 80 percent of the planet's surface is covered with flat plains that are probably solidified lava flows. These lava plains formed when lava flooded areas covering thousands of square kilometers and then solidified.

Most of the volcanoes on Venus are shield volcanoes. Shield volcanoes derive their name from their resemblance to ancient warriors' shields lying on the ground pointing upward. Shield volcanoes are often very large, but they have fairly gentle, rather than very steep, slopes. Shield volcanoes form when lava flows out from a single central vent. Rather than forming on the boundaries of tectonic plates, shield volcanoes usually form over a volcanic hotspot. These hotspots are places in the planet's crust where lava wells up from the planet's mantle. When the lava breaks through the surface, it erupts to form a shield volcano. Successive eruptions can form very large shield volcanoes.

Two of the larger known shield volcanoes on Venus are Sif Mons and Gula Mons. They have peak altitudes of about 4 kilometers above the surrounding surface, which compares to Mauna Loa, Earth's largest shield volcano, which rises about 8 to 9 kilometers above the Pacific ocean floor. (The largest shield volcano in the solar system is Mars's Olympus Mons, which towers about 25 kilometers above the Martian surface.) Often the top of a shield volcano will collapse to form a crater, known as a caldera. This collapse occurs when the lava flow retreats back to the planet's mantle, leaving nothing to support the top of the volcano. The calderas formed on Sif Mons and Gula Mons are about 100 kilometers across. Calderas are fairly common on the surfaces of both Venus and Earth. Calderas, however, are not the only types of craters found on Venus. Venus has many large impact craters that formed from meteorite impacts rather than volcanic activity.



Gula Mons, seen on the Venusian horizon (left), rises to an altitude of about 3 kilometers and is one of the larger shield volcanoes on Venus. Cunitz Crater can be seen in the center middleground. This image was returned by the Magellan spacecraft. (NASA/JPL)

The largest volcanic features found on Venus are coronae. These features are not found on the other terrestrial planets. Coronae, which are approximately circular in shape (hence their name, from the Latin for "crown"), form from an uplifting process. Hot mantle material swells and pushes the crust upward. Coronae usually have associated volcanoes and lava flows. Aine is a large corona on Venus that is about 300 kilometers in diameter. On a larger scale, Lakshmi Planum, which is part of Ishtar Terra—one of Venus's two large continental sized features—is about 1,500 kilometers at its widest point. Lakshmi Planum likely formed from the same process that formed the coronae, but on a larger scale.

Another common type of volcano found on Venus is the lava dome or pancake dome. Venus's lava domes are much smaller than its shield volcanoes, being typically tens of kilometers in diameter or less. They are usually circular and relatively flat—hence the name "pancake dome." They form when lava slowly flows out onto the surface and then flows back. However, a thin crust solidifies on the surface of the lava. When the lava subsides, the crust stays and cracks because it lacks support. Pancake domes are often found near coronae.

On Earth, volcanoes often form at the boundaries of the tectonic plates. Examples are the volcanoes on the western coasts of North and South America and the Mid-Atlantic Ridge, which runs along the Atlantic Ocean's floor between the North American and Eurasian plates. Such volcanoes are not found on Venus. Venus apparently does not have tectonic plates on its crust. Venus does not have plate tectonics similar to Earth's, but it does have tectonic activity. On Earth, plate tectonics is caused by convection currents in Earth's mantle slowly moving the crustal plates horizontally. On Venus, the crust is not divided into plates. Convection currents in the mantle cause vertical rather than horizontal crustal movement on Venus. Coronae are a good example of volcanic features formed on Venus from the crust's vertical tectonic motion.

Are the volcanoes on Venus still active, as on Earth, or are they extinct, as on Mars? Planetary scientists do not yet know the answer to this question. Despite being volcanically active, at any given time few of Earth's many volcanoes are actively erupting. The same would be true on Venus. Hence scientists would not expect to see volcanoes continually erupting on Venus, even if it is still volcanically active. Partly because of the planet's thick cloud layer, no one has observed a volcanic eruption on Venus, but there is some indirect circumstantial evidence to suggest that volcanoes on Venus are still active. Volcanic eruptions emit sulfur dioxide gas. Scientists observe frequent variations in the amount of sulfur dioxide in Venus's upper atmosphere. These variations could be caused by occasional volcanic eruptions spitting sulfur dioxide into the atmosphere. Space probes to Venus have also detected radio outbursts from Venus that are similar to those produced by lightning discharges from erupting volcanoes on Earth. These observations are evidence, but not proof, that Venus's volcanoes are still active. If planetary scientists were to observe a volcano on Venus in the act of erupting, then Venus would join Earth and Jupiter's satellite Io as the worlds in the solar system with still-active volcanoes.

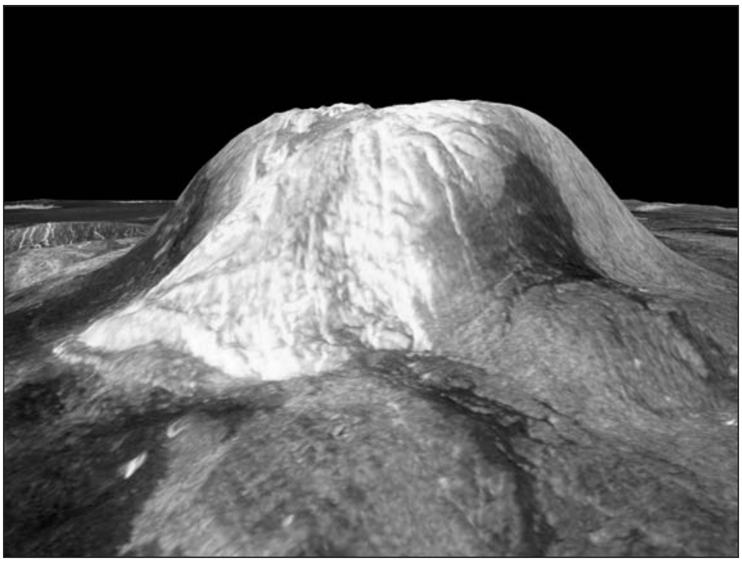
KNOWLEDGE GAINED

Because thick clouds veil the surface of Venus, astronomers for a long time could only speculate about the planet's surface. Prior to the space age, speculations varied: The surface was envisioned as hot and steamy by some, as a hot and dry desert by others. With the coming of the space age, astronomers were finally able to gather data on the surface of Venus. They did not, however, suspect just how hot Venus really was.

The first radar maps of Venus from Earth were made using the Arecibo radio telescope in Puerto Rico, beginning in the late 1970's. Because of the distance to the planet, these images had a relatively low resolution, on the order of a few kilometers. These Earth-based radar maps did, however, allow planetary scientists to observe the large-scale surface features of Venus.

Earth-based radar maps of Venus can reveal only part of the Venusian surface, because during Venus's closest approach to Earth only one side faces Earth. Orbiting spacecraft have therefore been sent to map the surface of Venus. Pioneer Venus 1 went into orbit around Venus in late 1978. This mission was the first orbital mission to use radar to map much of the surface of Venus, with a resolution of about 7 kilometers. In 1990, the Magellan mission went into orbit around Venus. The Magellan orbiter made much more extensive and detailed radar maps of Venus. Because it used a polar rather than an equatorial orbit, Magellan was able to map essentially the entire surface, including the polar regions, which were hidden to previous missions. The best resolution of the Magellan radar maps is about 100 meters.

The fact that Venus has volcanic activity, including both volcanic mountains and lava plains, in-



A close-up computer simulation of Gula Mons was generated from data returned by the Magellan spacecraft. It rises 3 kilometers above the area known as Eistla Regio. (NASA/JPL) dicates that, geologically speaking, the surface of Venus is very young. The surface, not the planet itself, is probably less than a half a billion years old.

CONTEXT

The volcanoes on Venus contribute to its very harsh surface environment. Volcanoes on Earth outgas significant amounts of carbon dioxide gas. Those on Venus and Mars probably do the same. On Earth, biological activity, such as plant respiration, uses the carbon dioxide. Earth therefore has a very small percentage of carbon dioxide in its atmosphere. On Venus, however, the carbon dioxide is still in the atmosphere; 97 percent of Venus's atmosphere is carbon dioxide. All this carbon dioxide produces a runaway greenhouse effect and surface temperatures greater than 700 kelvins.

Jupiter's moon Io is also volcanically active. However the volcanoes on Io differ from the volcanoes on Venus and other terrestrial planets. In addition to rock, Io's composition includes significant amounts of ice.

Venus, Mars, and Earth all have large volcanoes. However, the volcanic and tectonic activity is different on each of these three planets. The differences arise from differences in size and internal heating of the planets. Mars has had the least amount of volcanic activity. Earth's crust is divided into several tectonic plates. Movement of these plates is an important force in shaping the volcanic activity on Earth. Venus does not have a crust broken into several plates. Hence, Venus has much volcanic activity, but it does not have the types of features formed by crustal plate movement. Understanding how volcanic and tectonic activity differs among the various planets is one of the frontiers of planetary science.

Paul A. Heckert

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Chapter 9 of this very readable introductory astronomy textbook covers the planet Venus. There is a good section on the planet's surface volcanism.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. Chapter 11 of this introductory astronomy textbook is a complete and readable overview of Mercury, Venus, and Mars, including volcanic activity.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005. This textbook on the planets and satellites of the solar system summarizes our understanding of volcanic and other tectonic processes on Venus as well as other planets and moons.

Hester, Jeff, et al. *Twenty-first Century Astronomy*. New York: W. W. Norton, 2007. Chapters 6 and 7 of this well-illustrated astronomy textbook are about the terrestrial planets. Volcanic and tectonic processes are well covered, and the comparison of these processes on different planets is very good.

Morrison, David, Sidney Wolf, and Andrew Fraknoi. *Abell's Exploration of the Universe*. 7th ed. Philadelphia: Saunders College Publishing, 1995. Venus is covered in chapter 15 of this classic astronomy textbook.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. New York: Cambridge University Press, 2002. Chapter 9 of this astronomy textbook provides an overview of the terrestrial planets.

Zeilik, Michael, and Stephen A. Gregory. *Introductory Astronomy and Astrophysics*. 4th ed. Fort Worth, Tex.: Saunders College Publishing, 1998. Pitched at undergraduate physics or astronomy majors, this more advanced textbook goes into greater mathematical depth than most introductory astronomy texts. Chapters 4 and 5 cover the basic principles of the terrestrial planets, including volcanic processes.

See also: Earth's Atmosphere; Earth's Composition; Earth's Magnetic Field: Origins; Earth-Sun Relations; Extraterrestrial Life in the Solar System; Greenhouse Effect; Mars's Atmosphere; Planetary Atmospheres; Planetary Interiors; Planetary Magnetospheres; Planetary Rotation; Planetary Tectonics; Planetology: Comparative; Planetology: Venus, Earth, and Mars; Terrestrial Planets; Venus's Atmosphere; Venus's Craters; Venus's Surface Experiments; Venus's Surface Features.

W

White and Black Dwarfs

Category: The Stellar Context

White dwarfs are stars of about one solar mass in the last stage of their lives. They have no more ways to generate energy, so they shine only because they are very hot. As they radiate their energy away, they cool and fade, becoming cold, dark black dwarfs.

OVERVIEW

White dwarfs are a unique class of stars. They have high surface temperatures, at least initially, and low luminosities. This combination means that they are very small, about the size of Earth and thus much smaller than stars in the energy-generating stages of their lives. With a mass approximately the same as the Sun but contained in a sphere about the size of the Earth, they have average densities in the neighborhood of a billion kilograms per cubic meter.

The first white dwarf was discovered in 1862 by the American astronomer and telescope maker Alvan Graham Clark, while testing the 18.5-inch refracting telescope he made for Dearborn Observatory. He found that Sirius (Alpha Canis Majoris), the brightest appearing star in the night sky, had a very faint companion, given the name Sirius B. In 1896, a similar but even fainter companion was discovered orbiting the star Procyon (Alpha Canis Minoris) and was given the name Procyon B. Then Harvard astronomers discovered that the star 40 Eridani had a faint companion similar to Sirius B, which was named 40 Eridani B. Shortly thereafter, Adriaan van Maanen found a similar but still fainter single star (not a part of the binary system), which was subsequently named after him, Van Maanen's star. During the middle of the twentieth century, Willem Luyten of the University of Minnesota found several hundred of these white dwarf stars, as they had come to be called, by looking for faint blue stars with large proper motions. (The large proper motions showed that the stars were not very far away, so their faint appearance meant they really were intrinsically faint. The blue color showed they had hot surfaces, which, together with their faint luminosities, meant they were very small.) Today even more white dwarfs have been found with improved detectors and observing techniques.

By the early 1900's, it was known that these stars were intrinsically very faint, very small, and incredibly dense. Most, but not all, had a bluish-white color, indicating a hot surface. Modern observations have refined these early findings. Sirius B, the first white dwarf to be discovered, has a surface temperature of about 24,000 kelvins, a luminosity about 0.020 times the Sun's, a radius about 0.008 times the Sun's (about 5,600 kilometers, or about 0.9 times the Earth's radius), a mass about 1.1 times the Sun's, and an average density of about 3 billion kilograms per cubic meter (about 3 million times the density of water). The white dwarf 40 Eridani B has a surface temperature of about 12,000 kelvins, a luminosity about 0.004 times the Sun's, a radius about 0.014 times the Sun's (about 9,800 kilometers, or about 1.5 times the Earth's radius), a mass about 0.43 times the Sun's, and an average density of about 200 million kilograms per cubic meter (about 200,000 times the density of water).

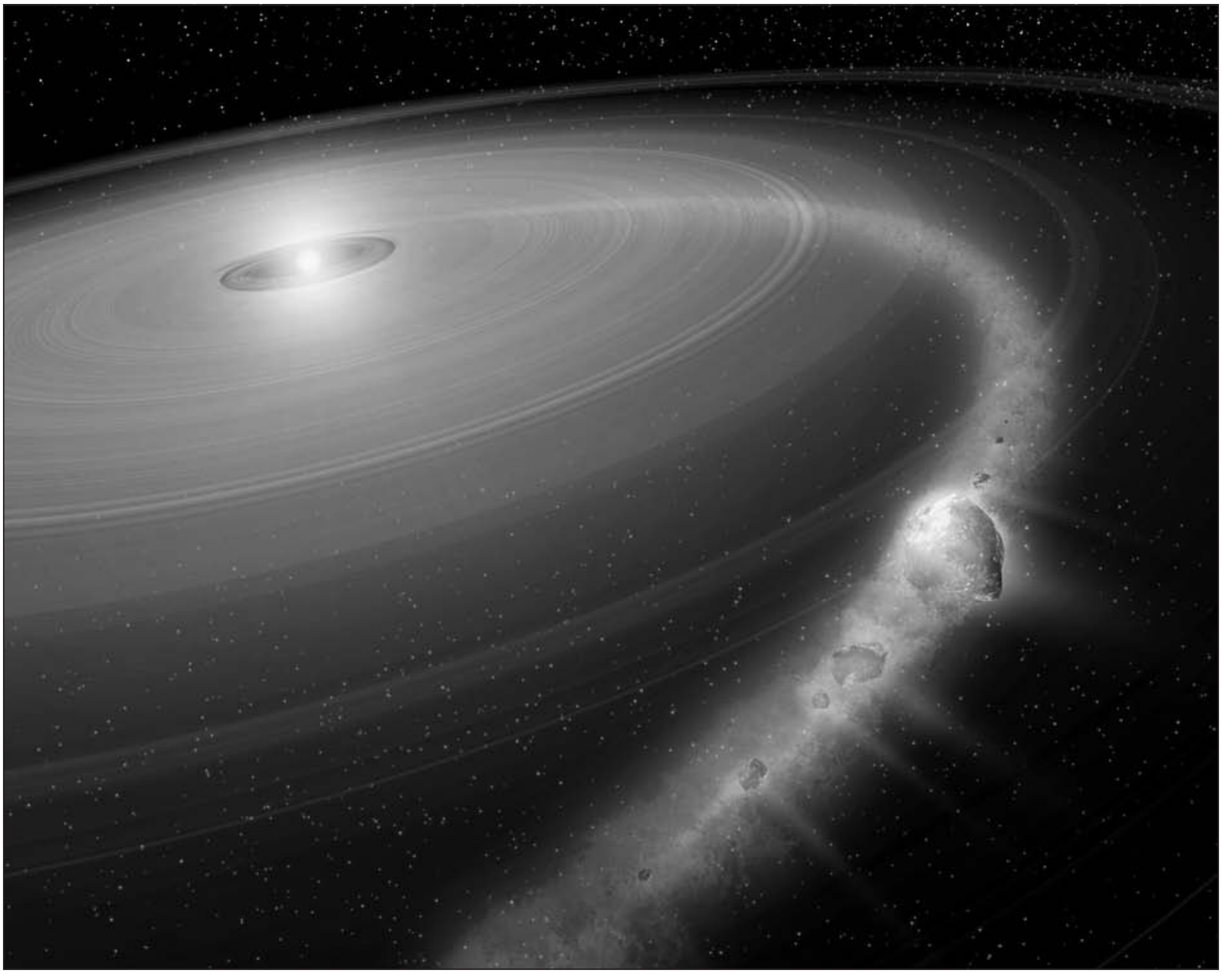
The unusual properties of white dwarfs, especially their great densities, initially made many astronomers question whether such stars could really exist or whether there was something wrong with the observations or the analysis of them. The English astrophysicist Sir Arthur Stanley Eddington summed up the scientific community's reactions this way:

> The message of the companion of Sirius, when decoded, ran: "I am composed of material three thousand times denser than anything you've ever come across. A ton of my material would be a little nugget you could put in a matchbox."

> What reply could one make to something like that? Well, the reply most of us made in 1914 was, "Shut up; don't talk nonsense."

The theoretical work to try to figure out this "nonsense" began with Eddington himself, who in 1924 suggested that the ultrahigh density of white dwarfs might be caused by their extremely high temperatures, which would totally ionize of the matter in them. The resulting bare nuclei and free electrons could be forced into a much smaller volume of space than un-ionized atoms surrounded by orbiting electrons could be packed.

In 1927, Ralph Howard Fowler used the newly developed Fermi-Dirac statistics (named for Enrico Fermi and Paul Adrien Maurice Dirac), and the Pauli exclusion principle (named for Wolfgang Pauli) from quantum mechanics to show that at the extreme densities within white dwarf stars, electrons are completely degenerate. This means that the electrons fill all the "cells" in combined momentum-position phase space up to some maximum momentum, with two and only two electrons in each "cell." Unlike an ideal gas, which releases gravitational energy when it contracts, a sphere of degenerate electrons cannot shrink without an input of energy to increase the momentum of the electrons; shrinking reduces the position part of phase space, and this requires an increase in the momentum part of phase space. Consequently, without an input



An artist's rendition of a comet (right) as it is broken apart by the gravitational field of white dwarf G29 38 (upper left). (NASA/JPL-Caltech/T. Pyle, SSC) of energy, degenerate electrons behave collectively like an incompressible fluid.

With high enough densities, the upper momentum states of degenerate electrons are relativistic; that is, the fastest electrons move at a speed that is a significant fraction of the speed of light. In the early 1930's, Subrahmanyan Chandrasekhar applied relativity theory to derive an equation of state for relativistically degenerate electrons. The pressure of the degenerate electrons far exceeds the pressure of the nuclei, which still behave like an ideal gas. Thus it is degenerate electron pressure that balances gravity to maintain hydrostactic equilibrium in white dwarfs. Employing the degenerate electron equation of state and the condition of hydrostatic equilibrium, Chandrasekhar was able to compute models for the interior structure of white dwarfs of various masses. He found that the greater is the mass of a white dwarf, the smaller its size is and the greater its density is. This is shown in the data for Sirius B and 40 Eridani B: Sirius B is more massive, but smaller and denser, than than 40 Eridani B. Chndrasekhar's model implied that there is an upper mass limit for white dwarf stars, now called the Chandrasekhar limit, of about 1.4 solar masses. Although only a few white dwarfs in binary systems have had their masses measured directly, they are below and thus consistent with the Chandrasekhar upper mass limit. For his wide-ranging contributions to astrophysics throughout his career, of which his work on the physics of white dwarfs was just the beginning, Chandrasekhar was awarded the Nobel Prize in Physics in 1983.

The spectra of white dwarf stars are difficult to interpret. Their high surface gravity leads to pressure broadening of the absorption lines in their spectra, and in some cases (such as the white dwarf Wolf 489) the lines become so broad and shallow that they almost disappear into the continuous background spectrum. Some white-dwarf spectra display prominent hydrogen lines, others show strong helium lines, and still others have lines representing a mixture of hydrogen, magnesium, potassium, calcium, iron, or a combination of these elements. White dwarf spectra also show a very large Stark effect (named for Johannes Stark), caused by the presence of strong electrostatic fields that result from high densities and the accompanying degeneracy. The spectra of some white dwarfs indicate the presence of magnetic fields up to about 10 million times stronger than that of Earth.

APPLICATIONS

Eddington realized that the large density and high surface gravity of white dwarfs provided a testing ground for a prediction of Albert Einstein's general theory of relativity. According to general relativity, light emitted in the presence of a strong gravitational field (such as exists at the surface of a white dwarf) is redshifted; basically, photons of light lose energy climbing out of the strong gravitational field, thus lengthening their associated wavelengths. At Eddington's request, Walter Adams at Mount Wilson Observatory attempted to measure this gravitational redshift in white dwarf spectra. His measurements were not accurate enough for a definitive confirmation, though they were consistent with the prediction of general relativity.

In 1954, using better instrumentation, D. M. Popper measured the wavelengths of the lines in the spectrum of 40 Eridani B with greater accuracy. He found that, after allowing for the Doppler shift of the spectral lines due to the star's radial velocity, there was a residual wavelength increase of 0.0070 percent. Based on the mass and radius of this star, general relativity predicts its spectral lines should have their wavelengths increased by 0.0057 percent. (These percentage shifts mean that for a spectral line with a "normal" wavelength of 500 nanometers, general relativity predicts its wavelength should be increased by 0.029 nanometer, while Popper's measurements corresponded to an increase of 0.035 nanometer—very good agreement considering the difficulty in accurately measuring the wavelengths of white dwarf spectral lines.) The confirmation of a prediction of general relativity through observation of a gravitational redshift in a white dwarf spectrum was a major achievement.

CONTEXT

All stars spend most of their energy-producing lives as main sequence stars, fusing hydrogen into helium in their cores. White dwarfs and black dwarfs now are understood as stars in the last stages in their lives, with less than about 8 solar masses when first formed. When stars with initial masses below about 8 solar masses exhaust the hydrogen in their cores, they expand to become red giants, eventually fusing helium into carbon and maybe oxygen in their cores. However, they are not massive enough to generate energy by any other nuclear fusion reactions. Strong stellar winds and thermal pulsations in their bloated atmospheres puff off their outer layers as expanding shells of gas called planetary nebulae (a term that has nothing to do with planets but, rather, originated in the 1800's when, with the telescopes available then, these objects looked round, like planets, and fuzzy, like nebulae). The stars remaining at the centers of planetary nebulae are the former cores of red giants, exposed to view as the planetary nebulae expand and dissipate. These central stars of planetary nebulae are progenitors of white dwarfs.

To become white dwarfs, stars must lose enough of their original mass during the last stages of their lives—whether by strong stellar winds, planetary nebulae, or some other mechanism—that their final mass is less than the Chandrasekhar limit of 1.4 solar masses. Stars with initial masses below about 0.25 to 0.5 of a solar mass are not massive enough to ignite helium fusion in their cores and probably will never become red giants or planetary nebulae; instead they may progress slowly from the main sequence directly to the white dwarf stage, taking perhaps hundreds of billions to trillions of years to do so.

White dwarf stars can no longer generate energy through any nuclear fusion process. They cannot release gravitational energy by contracting because their electrons are degenerate. They shine only because they are very hot, with central temperatures perhaps as high as 100 million kelvins initially. As they shine, they radiate their energy away, slowly cooling and fading, like an ember plucked from a fire. At the initial high temperatures, the atomic nuclei in white dwarfs behave like an ideal gas, but as the stars cool, the nuclei "freeze" into a regular lattice-like pattern, similar to a giant crystal, through which the degenerate electrons move freely. The white dwarfs, now essentially solid, continue to cool and grow fainter, eventually becoming cold, dark black dwarfs. This is the fate that awaits our Sun several billion years in the future.

However, this may not be the end for all white dwarfs. If a white dwarf is a member of a close binary system with a red giant companion, gas (mostly hydrogen) can be transferred from the red giant onto the white dwarf. The hydrogen that accumulates on the white dwarf's surface may be heated sufficiently to fuse explosively into helium. A shell of hot gas is blasted into space, becoming as much as 100,000 times brighter than the white dwarf itself. This explosive outburst is called a nova. Since the process can repeat over and over again, novae can recur for decades or even centuries.

If the white dwarf in a close binary system is already almost at the Chandrasekhar limit (the maximum mass a white dwarf can have), any additional matter transferred from the companion can push the white dwarf over the mass limit. If this happens, the white dwarf collapses on itself and heats up to about a billion kelvins. The high temperature initiates a series of nuclear fusion reactions that blow the star apart as a Type Ia supernova. (A Type II supernova, on the other hand, is produced by a massive supergiant that explodes once it develops an iron core that collapses.) Because Type Ia supernovae all are produced in the same way by essentially identical objects (white dwarfs gaining enough mass to exceed the Chandrasekhar limit), they reach approximately the same peak luminosity, nearly 10 billion solar luminosities, making them reliable "standard candles." Since their peak luminosity is so high, they can be used as standard candles to determine the distances of galaxies billions of light-years away.

V. L. Madhyastha

FURTHER READING

Chaisson, Eric, and Steve McMillan. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008. Very well-written college-level textbook for introductory astronomy courses. Several chapters deal with the later stages of stellar evolution, including white dwarfs, black dwarfs, recurrent novae, and Type Ia supernovae.

Chandrasekhar, Subrahmanyan. "On Stars, Their Evolution, and Their Stability." *Reviews of Modern Physics* 56 (1984): 137-147. Although this is a technical article, it is seminal: the Nobel lecture delivered by Chandrasekhar in 1983. The general reader may easily skip over the mathematical equations and concentrate on the text, which is highly informative and easy to comprehend. Includes references to several important works on the subject.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006. A well-written, thorough college textbook for introductory astronomy courses. Covers not only white and black dwarfs but also the lifetimes of stars generally, including recurrent novae, and Type Ia supernovae.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory astronomy textbook. Addresses the various paths of star evolution, including white dwarfs.

Kafatos, Minas C., and Andrew G. Michalitsianos. "Symbiotic Stars." *Scientific American* 251 (July, 1984): 89-94. An informative article that describes some of the intriguing aspects of binary systems in which matter is transferred from a red giant to a hot white dwarf.

Luyten, Willem J. "White Dwarfs." In *Advances in Astronomy and Astrophysics*. Vol. 2. New York: Academic Press, 1963. Luyten succinctly introduces all the observed properties of white dwarfs and major problems facing astronomers. Having spent his career identifying and studying the strange and unusual features of these stars, he presents a readable introduction with a list of some thirty-seven white dwarfs and relevant references.

Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008. Very thorough college textbook for introductory astronomy courses. Divided many short sections on specific topics, with several that address the later stages of stellar evolution, including white dwarfs, black dwarfs, recurrent novae, and Type Ia supernovae.

Van Horn, Hugh M. "Physics of White Dwarfs." *Physics Today* 32, no. 1 (1979): 23. An elementary introduction to the mathematically involved subject of what occurs inside a white dwarf. Van Horn provides a list of useful references.

Weidemann, V. "Masses and Evolutionary Status of White Dwarfs and Their Progenitors." *Annual Review of Astronomy and Astrophysics* 28 (September, 1990): 103-137. A lengthy article with numerous references describing developments in observing and modeling the evolution of white dwarfs. The same issue contains an article on the cooling of white dwarfs by Francesca D'Antona and I. Mazzitelli. The articles, although technical, are readable by nonspecialists.

See also: Brown Dwarfs; Gamma-Ray Bursters; Gravity Measurement; Hertzsprung-Russell Diagram; Main Sequence Stars; Novae, Bursters, and X-ray Sources; Nuclear Synthesis in Stars; Protostars; Pulsars; Red Dwarf Stars; Red Giant Stars; Solar Evolution; Stellar Evolution; Supernovae; Thermonuclear Reactions in Stars.

X

X-Ray and Gamma-Ray Astronomy

Category: Scientific Methods

X-ray and gamma-ray astronomy involves the observation of events in the universe that occur at energies far greater than what is normally shown by visible light or other forms of astronomy. This branch of astronomy promises great advances in our understanding the formation and fate of matter in the universe.

OVERVIEW

Like visible light, both X rays and gamma rays are electromagnetic radiation, but at much shorter wavelengths (and correspondingly higher frequencies). Because of this extreme difference, it is much more difficult to collect large quantities of X rays and gamma rays and to detect a source. In general, the field is called high-energy astrophysics, a name that also refers to cosmic rays (atomic and subatomic particles spewed by various nuclear reactions).

Because of the extremely short wavelengths involved, X rays and gamma rays are both measured in terms of their energy, from thousands to billions of electron volts (eV). Although there is no firm demarcation, X rays span the range of 1,000 to 100,000 eV and gamma rays from 100,000 eV upward (visible light is around 1-2 eV). X rays are emitted by the innermost electrons of an atom releasing energy or through Bremsstrahlung (braking) radiation. Gamma rays, by definition, are emitted by reactions in the nuclei of atoms or by subatomic reactions. There is some overlap between the two spectral bands, and it may be impossible to tell if an emission was from an energetic electron or a nuclear reaction.

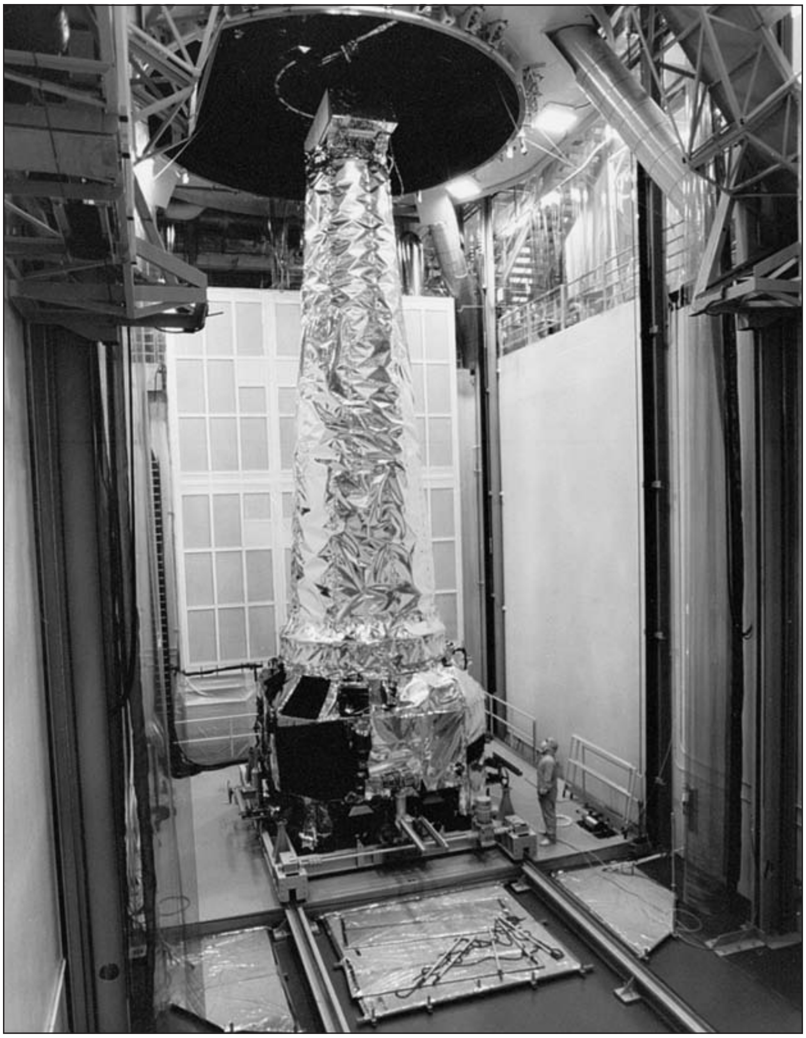
Since X rays and gamma rays both pass through solid matter, it is difficult to focus them onto a detector and difficult to make them interact with the matter in the detector itself. A common method of "focusing" high-energy radiation is to use collimators to exclude all radiation except that coming from a particular direction. The basic technique can be imitated by looking through a cluster of straws held at arm's length: The view is narrow and restricted. Essentially, all light that is nearly "on axis" (that is, parallel with the centerline of the straws) passes through, while light that is "off axis" strikes the sides of the straw. The longer the straws, the narrower the field of view. A variety of complex collimation techniques have been developed to allow only radiation from a desired source to fall directly on a detector. While simple, this method generally yields images of the sky that lack the finer resolution of optical telescopes and thus leaves much uncertainty about the location of energy sources.

The most direct solution to this problem is the use of grazing-incidence mirrors to focus "soft" (low energy) X rays. Grazing incidence occurs when radiation strikes a surface at an extremely shallow angle and is reflected rather than being absorbed or scattered. The effect is readily seen when visible light strikes a windshield or pond at less than the critical angle and is reflected to cause glare. Because of the energy of X rays, the angle of incidence must be extremely shallow, typically less than 1°, and the surface must be exceptionally smooth to allow an image to form. X-ray telescopes generally use two reflectors—a parabolic primary and a hyperbolic secondary—to focus the radiation into an image without aberration. To provide the shallow angle of incidence, the mirrors resemble tubes (the segment of the paraboloid or hyperboloid surface is more distant from the focus than with conventional telescopes). This also requires that the secondary mirror be mounted directly behind and precisely aligned with the primary. Such an arrangement is known as a Wolter Type 1 telescope. A number of variations are available. Because only a small region of the radiation is intercepted by the mirrors, modern X-ray telescopes typically will use a nesting scheme in which up to six complete telescopes are built within each other.

Another type of X-ray telescope is the Kirkpatrick-Baez, which uses curved plates of glass in banks one behind the other. This arrangement will focus the light in one axis, then the other. While lacking the fine resolution of Wolter telescopes, the Kirkpatrick-Baez arrangement is well suited to all-sky surveys.

Grazing-incidence telescopes were developed in the 1960's and 1970's. The 1980's saw the development of a radically new approach: normal-incidence X-ray mirrors. In Bragg crystals, the internal structure of a crystal can refract X rays almost as effectively as glass refracts light. Normal-incidence mirrors are built up in layers of microscopic crystals that intercept the X rays and reverse their direction. When laid upon parabolic or hyperbolic surfaces, one can build an X-ray telescope that resembles a conventional reflector telescope and provide comparable resolution for low- to medium-energy X rays.

At higher energies, though, the efficiency of reflectors decreases until virtually no image is being focused. At very high energies, a variation on the pinhole camera technique may be used. In this approach, the front end of the telescope is a plate punctured by holes in a quasi-random pattern. Each pinhole allows an X-ray image to be projected onto the detectors behind the plate. However, because so many images are projected at once, the scene must be mathematically deconvolved (decoded) to produce a true picture of the scene.



The Chandra X-Ray Observatory undergoing inspection at TRW Space and Electronics Group, Redondo Beach, California, in 1998. (NASA/TRW)

There is no way to focus gamma rays. Their energies are simply so high that they would pass unhindered through the mirrors. Instead, gamma-ray "telescopes" constitute instruments that are designed to detect interactions of gamma rays with matter.

X rays and gamma rays, because of the energies involved, are detected less directly than light; specifically, the energy yielded by interaction with some intermediate object is detected. The first and simplest X-ray detector is photographic film. Exposure to X rays will cause the film to darken (that is, produce a negative image). Film is used only in X-ray telescopes that can be recovered, such as crewed satellites and sounding rockets;

various electronic detectors are used in uncrewed satellites. The two are complementary in use. Film provides the highest spatial resolution and reveals fine details of the source and its position relative to other objects, but over a broad energy range. Electronic detectors can measure energy levels with a fair degree of accuracy but at the expense of spatial resolution. In both cases, the energy resolution can be enhanced by placing filters in front of the detector, but this generally establishes a low-energy cutoff and higher-energy radiation will still pass through.

Most development work has gone into electronic detectors that measure the energies yielded when an X-ray or gamma-ray photon strikes matter and releases enough electrons to generate a current, or stimulates electrons to release lower-energy photons as visible light, which can then be detected by conventional devices.

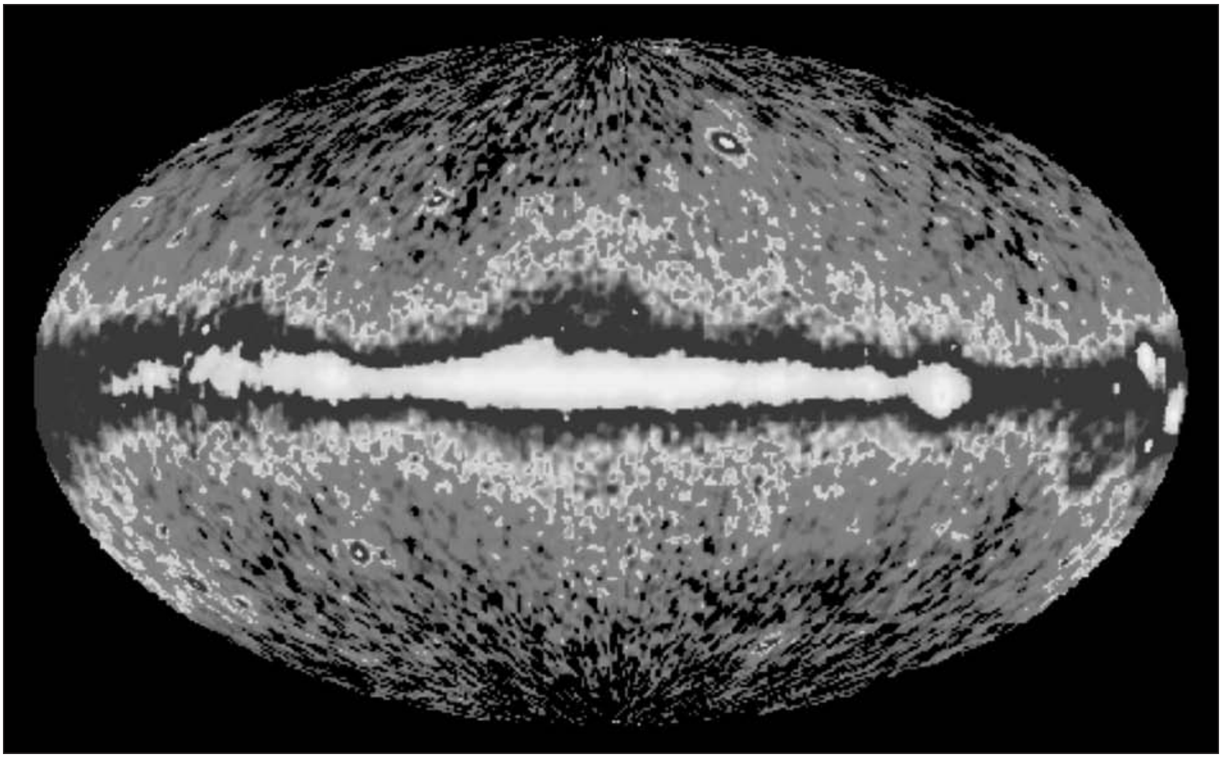
Proportional counters are similar to Geiger-tube counters. A high-energy photon enters a gas-filled chamber, intercepts a gas atom or molecule, and generates an ion and an electron. These are attracted to the electrodes (anode and cathode) in the tube and cause an increase in the electrical current that is measured by the instrument's electronics. In its simplest form, the counter simply registers the arrival of a photon regardless of direction. Most proportional counters, though, are actually gangs of many small counters mounted behind collimators to narrow the field of view. They also use shielding to reduce the chance of photons entering from behind the detector and giving a false signal. In some cases, the collimators move so as to make sources wink on and off and thus improve the precision of its location and even to deduce the structure of extended sources.

To increase the chances of detecting an X ray, detectors sometimes are electrically charged so that a single X ray will produce a shower of electrons that is readily detected (comparable to electric eyes that generate an electric current when struck by light). The electrically charged device is a microchannel plate, which is like a microscopic collimator with its channels as a slant. This ensures that an X ray will strike the wall of the channel, which has an electric charge not quite high enough to cause an electric arc. The incoming X ray provides the needed extra jolt and releases a cascade of electrons that exit the back of the plate to be measured by the wires of a grid behind the detector.

In the late 1980's, progress was made on developing X-ray charge-coupled devices (CCDs) that could provide a more efficient means of creating X-ray images. CCDs are electronic analogs of the retina and comprise thousands of small photosensors that are read individually by computer. Unlike conventional CCDs, an X-ray CCD will be illuminated from behind, again so that the incident X rays will generate light or an electrical charge that is detected by the CCD.

Spectral measurements in X rays use proportional counters, too, but require a grating or crystal to spread the X rays. The familiar visible-light analog is a wedge prism that breaks white light into the color spectrum. With X rays, a similar effect is desired since the intensity of the X rays at different energies is a measure of the activity at the source. Gratings are surfaces that are microscopically ribbed. The degree to which radiation scatters depends on the energy of the photon. A proportional counter can then be moved along an arc and radiation can be measured. Spectral and imaging instruments can be combined by placing a transmission grating or crystal in the telescope's line of sight. The focused radiation then is spread across the imager and its intensity measured. Normally, this technique works only for strong X-ray sources.

In scintillation detectors, the incoming radiation is trapped in a crystal and causes a flash of light, which is measured by a photomultiplier tube. The most sensitive of these instruments require that the crystal be supercooled by liquefied or solid gas so that the body temperature of the crystal does not cause a false reading. Such crystals can be used in spectrometers. A variation of the effect is Compton scattering, in which a high-energy X ray strikes matter and is scattered at lower energy and with release of an electron whose energy corresponds to the original energy. At higher energies, however, the efficiency of this effect declines. Gamma rays can also produce electron-positron (antielectron) pairs that can be detected in spark chambers, which cause sparks between electrified plates



This image, made possible by the Compton Gamma Ray Observatory, shows the entire sky in the gamma range above 100 million electron volts. Without an orbiting observatory such as Compton, such an image would not be possible to acquire from Earth, since the atmosphere protects the planet's surface, and hence terrestrial life, from the penetration of gamma radiation. The bright middle band is the gamma radiation from the Milky Way. (NASA, Compton Gamma Ray Observatory)

or wires. These are then registered by photomultiplier tubes.

Because gamma rays cannot be focused, different techniques are used to determine the shape of a source. The instrument can be tilted back and forth and the flux change measured to determine its origin. Or, the instrument can be built as a pair, one behind the other, and the two sets of signals collated to determine the origin of each gamma ray.

X-ray and gamma-ray detectors, in general, have anticoincidence detectors that are designed to detect cosmic rays and thus allow the signals they generate to be subtracted from the instrument signal, somewhat like filtering the noise in a radio.

METHODS OF STUDY

X-ray and gamma-ray astronomy is conducted above the atmosphere of the Earth, since the atmosphere absorbs all X rays (even the thin layer above 36,000 meters absorbs soft X rays). While both fields are best conducted from satellites, both started on (and still use) suborbital platforms: X-ray astronomy from suborbital rockets and gamma-ray astronomy from balloons because the lower fluxes required larger, heavier detectors. Suborbital rockets (also called sounding rockets) can expose a payload to the space environment for several minutes, depending on the weight of the instrument and the power of the rocket. Suborbital rockets were a major tool during much of early X-ray astronomy and continue to serve the same purpose as new telescopes are developed and tested. They also have filled the gap between flights of major X-ray satellites. Gamma-ray instruments, though, have to be larger and spend more time at altitude to intercept a sufficient flux from gamma-ray sources, so uncrewed balloons have been their preferred suborbital platform. In either case, satellites are the best means of operation, since they provide essentially indefinite observing time for medium to heavy instruments.

Several X-ray astronomy satellites have been launched, the most notable being the High-Energy Astrophysical Observatories (HEAO), the European Space Agency's X-ray satellite EXOSAT, and the Roetgen X-Ray Satellite (ROSAT). Three of these satellites—HEAO 2, EXOSAT, and ROSAT—carried Wolter Type 1 X-ray telescopes to produce images and spectra of the heavens.

Between 1967 and 1969 Orbiting Solar Observatory (OSO) 3 instrumentation registered 621 gamma-ray photons, confirming a diffuse gamma-ray background, but it must be noted that bursts as discovered by the military Vela satellites remained classified until 1973 rather than immediately becoming a subject of scientific investigation.

In 1972, Small Astronomy Satellite (SAS) 2 discovered an unexpected point source of gamma rays later identified as the neutron star Geminga. As a result of the Cosmic Origins Spectrograph (COS) B observations, the number of gamma-ray point sources was raised to twenty-five between 1975 and 1981. The first extragalactic gamma-ray source was identified as the quasar 3C 273. The study of gamma-ray astronomy greatly expanded during the two years (1979-1981) of HEAO 3 observations. One of its most important detections was 511-kilo-electron-volt radiation coming from the annihilation of electrons and positrons in the center of the Milky Way. The long-lived (1980-1989) Solar Maximum Mission observatory discovered soft gamma rays originating from solar flares.

Gamma-ray astronomy in the 1990's was conducted primarily by the National Aeronautics and Space Administration's (NASA's) Compton Gamma Ray Observatory and the European Space Agency's (ESA's) BeppoSax. However, ground-based Atmospheric Cherenkov Telescopes (ACTs) detected hard gamma radiation coming from blazars; emissions varied on timescales ranging from a few minutes to several hours.

NASA's High Energy Transient Experiment (HETE) 2 and the Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) satellites and ESA's International Gamma-Ray Astrophysics Laboratory (INTEGRAL) continued to expand the nature of high-energy astrophysics research until NASA launched the Swift satellite in 2004, which began finding as many as one hundred gamma-ray bursters annually, and doing so in such a way that the source of the intense gamma-ray emission could be identified quickly enough for other observatories, such as Chandra, to begin rapidly recording the afterglow of the bursters. Swift data provided evidence for mergers of neutron stars or the collapse of a neutron star into a black hole. When NASA launched the Gamma-Ray Large Area Space Telescope on June 11, 2008, gamma-ray astronomers anticipated another tremendous advance in their discipline.

The principal X-ray observatory for the twenty-first century is the Advanced X-Ray Astrophysics Facility (AXAF), conceived as an X-ray complement to the optical Hubble Space Telescope. Once in orbit after delivery by the space shuttle *Columbia* mission STS-93 in July, 1999, AXAF was renamed the Chandra X-Ray Observatory, or CXO. Chandra has a set of six nested Wolter Type 1 mirror pairs (each pair equivalent to a single telescope) to produce X-ray images with resolutions of 0.1 arc second or better. It carries advanced instrumentation derived from the HEAO 2 suite of instruments to produce images and spectra of X-ray objects. Also included is a unique X-ray CCD camera that takes advantage of the latest in solid-state detectors.

The Compton Gamma Ray Observatory (GRO) was launched on the space shuttle *Atlantis* in 1991. It carried four different but complementary gamma-ray instruments. Because gamma rays cannot be focused, they are not arranged in a common focal plane, like detectors in the Hubble Space Telescope, but stand alone on the satellite bus. GRO's instruments spanned a broad energy range, from 10,000 to 10 billion eV. The Burst and Transient Source Experiment (BATSE) recorded gamma-ray "flashes" that have puzzled scientists since 1967. These flashes may be caused by matter falling into neutron stars or black holes, but their unpredictability has hampered efforts to locate them. The burst experiment involved eight detectors that surveyed the entire sky to record and locate bursts so that optical telescopes and other types of telescopes could identify candidates for the burster. The Oriented Scintillation Spectrometer Experiment (OSSE) measured the spectra of objects when gamma rays caused scintillations in crystals. The four scintillation units were tilted to observe "empty" sky next to a source and subtract its background noise in order to obtain the signal of the object being observed. The Imaging Compton Telescope (COMPTEL) was actually two detector arrays using the Compton effect. Light was scattered and recorded as the gamma-ray encounters a scintillation liquid in the first detector, and the gamma ray was reemitted at lower energy to be detected in the same manner at the next level. The energy of the two emitted photons revealed the energy of the original gamma ray, and the locations of the two "hits" pointed back toward the source.

GRO's Energetic Gamma Ray Experiment Telescope (EGRET) was a large instrument designed to detect even low fluxes of the highest-energy gamma rays. Gamma rays entering the telescope struck a tantalum sheet and created an electron-positron pair that then traveled through two spark chambers and finally into a crystal scintillator. A complex anticoincidence system discounted sparks caused by cosmic rays. The energy of the gamma ray was recorded and its direction was revealed by the paths of the two particles it creates.

GRO succeeded in detecting gamma-ray bursters. It has also found antimatter clouds above the Milky Way and uncovered a bizarre new type of star, a "bursting pulsar," that flashes gamma rays. GRO's mission concluded on June 4, 2000, when the massive observatory was directed to plunge through the upper atmosphere in such a way that debris was not likely to fall on populated areas. There were safety concerns about large pieces of debris that might well survive reentry if GRO fell out of orbit in an uncontrolled fashion; however, there were no confirmed reports of debris making it to the deserted South Pacific Ocean area chosen for impact.

CONTEXT

X-ray and gamma-ray astrophysics has become one of the most revealing disciplines in astrophysics since its development following World War II. Extraterrestrial cosmic rays and X rays were detected in the 1800's by balloon crews who carried electrostatic instruments and cloud chambers aloft. Even when their true meaning became known, it was not appreciated that the sources might be caused by stars. As laboratory physics developed an understanding of nuclear fusion and how matter decays, it became obvious that X rays and gamma rays were generated by stars. Nevertheless, it was doubted that the flux (or total energy flow) would be great enough to be measured, at least for any star other than Earth's Sun.

With the availability after World War II of captured V-2 rockets to carry instruments aloft in tests, however, some rudimentary instruments were flown, followed by larger instruments aboard uncrewed balloons starting in the 1950's. It was discovered that as sensitivity increased, what could be seen became richer and more detailed. In the early 1960's, it appeared that the sky was suffused with a strong background glow of X rays. The Uhuru (Swahili for "freedom") satellite carried an all-sky survey detector that had sufficient resolution to detect more than three hundred discrete sources among the background glow. This led to the development and launch of a series of three High-Energy Astrophysical Observatories. HEAO 1 and 3 carried detectors that mapped the entire sky in X rays through gamma rays (HEAO 3 also carried cosmic-ray detectors). HEAO 2 carried the first stellar X-ray telescope and discovered that the X-ray background was composed largely of point sources that could not be distinguished at lower resolutions. This led to the discovery of the X-ray components of known visible objects and of previously unknown objects. In many cases, what was seen in X rays matched very nicely with the visible and radio components. In other cases, it seemed as though two different objects were being viewed.

Comparable work has been done in solar physics. The Orbiting Solar Observatories (OSOs) carried a number of X-ray and gamma-ray instruments in the 1960's and 1970's.

The Skylab space station in 1973-1974 included X-ray spectrometers and imaging telescopes in its array of eight solar telescopes. The Solar Maximum Mission (SMM) satellite (1980-1989) carried X-ray burst detectors and gamma-ray spectrometers to measure the output of the Sun.

In general, astronomers have found that the universe is far more energetic than previously believed. The Crab nebula—the remnant of a star that exploded in 1054—was found to pulse in X rays at the same rate as its visible pulsar, thus suggesting a very compact object. X rays have been found coming from the cores of quasars and most "normal" galaxies. The heart of the Milky Way seems to be the site of matter-antimatter annihilation. Instruments have measured X rays being emitted at an energy of 511,000 eV (or 5.11 keV), which corresponds to the conversion of electrons and positrons (antielectrons) into energy.

Gamma-ray astronomy provides a different view of the universe. Specifically, it reveals the creation and destruction of matter (properly, its conversion from energy to matter and back) in supernovae, neutron stars, pulsars, black holes, quasars, active galaxies, and other objects. Gamma-ray astronomy has confirmed that the heavier elements are created in the blast furnace of a supernova when a star self-destructs. The famous Supernova 1987A produced gamma-ray lines indicating that cobalt 56 was decaying into iron. Cobalt 56 is unstable and had to have been created shortly before the observation—that is, when the star exploded.

The most important result from high-energy astronomy—as well as from radio and infrared—is the increasing awareness that objects must be studied in terms of their total output rather than as emitters in different spectral bands. What puzzles scientists in one band may be solved in another, or at least a new line of investigation can be illuminated.

Dave Dooling

FURTHER READING

Arny, Thomas T. *Explorations: An Introduction to Astronomy*. 3d ed. New York: McGraw-Hill, 2003. A general astronomy text for the nonscience reader. Includes an interactive CD-ROM and is updated with a Web site. Comets are covered.

Fabian, A. C., K. A. Pounds, and R. D. Blandford. *Frontiers of X-Ray Astronomy*. New York: Cambridge University Press, 2004. For the most serious astronomy reader or students of astrophysics. Covers contemporary research with space-based X-ray telescopes.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2008. College-level introductory text covering the field of astronomy. Contains descriptions of astrophysical questions and their relationships. Informative.

Hirsh, Richard F. *Glimpsing an Invisible Universe: The Emergence of X-Ray Astronomy*. New York: Cambridge University Press, 1983. History of X-ray astronomy through the High Energy Astrophysical Observatories. Written for the well-informed reader.

Maccarone, Thomas J. *From X-Ray Binaries to Quasars: Black Holes on All Mass Scales*. New York: Kindle, 2006. Provides descriptions of high-energy processes that produce X-ray emissions. Covers the cosmological significance of quasars, black holes, and other high-energy objects.

McLean, Ian S. *Electronic and Computer-Aided Astronomy: From Eyes to Electronic Sensors*. Chichester, England: Horwood, 1989. A survey of the history and applications of electronics and astronomy. While much of the book is a technical survey, several chapters provide a general history and explanation of CCDs.

Schlegel, Eric M. *The Restless Universe: Understanding X-Ray Astronomy in the Age of Chandra and Newton*. New York: Oxford University Press, 2002. Covers X-ray astronomy for early rocket-launched studies to the Chandra and XMM-Newton observatories. Explains the cosmological implications of current X-ray astronomy research. For the serious layperson.

Trumper, Joachin, and Gunther Hasinger. *The Universe in X Rays*. New York: Springer, 2008. Reviews the development of X-ray astronomy and its impact upon advancements in astrophysics and cosmology. Covers

ROSAT, RXTE, BeppoSax, Chandra, and XMM-Newton results.

Tucker, Wallace. *The Star Splitters: The High Energy Astronomy Observatories*. NASA SP-466. Washington, D.C.: Government Printing Office, 1984. History of the HEAO satellite program and an overview of its results. Includes descriptions of X-ray optics and detectors.

Verschuur, Gerrit L. *The Invisible Universe: The Story of Radio Astronomy*. New York: Springer Praxis, 2006. Provides a history of developments in radio astronomy, and along the way describes the discovery of pulsars, quasars, and radio galaxies. Suitable for a general science course in college as well as for astronomy majors as background information.

Weeks, T. C. *Very High Energy Gamma Ray Astronomy*. New York: Taylor & Francis, 2003. Covers gamma-ray astronomy through results from the Compton Gamma Ray Observatory. For students of either theoretical or experimental high-energy astrophysics.

See also: Archaeoastronomy; Coordinate Systems; Earth System Science; Gamma-Ray Bursters; Gravity Measurement; Hertzsprung-Russell Diagram; Infrared Astronomy; Neutrino Astronomy; Optical Astronomy; Radio Astronomy; Telescopes: Ground-Based; Telescopes: Space-Based; Ultraviolet Astronomy.

Appendixes

Glossary

Absolute magnitude: The brightness of a star or other celestial body measured at a standard distance of 10 parsecs. *See also* Apparent magnitude, Luminosity, Parsec.

Absolute space, absolute time: An absolute idea that there is an overlying stationary structure in space and time which never changes, and against which all events and objects can be measured. The idea now generally is considered outdated in terms of relativity theory.

Absolute temperature scale: A temperature scale which sets the lowest possible temperature (absolute zero, or the temperature at which molecular and atomic translational motion stops) at zero. *See also* Kelvin.

Absolute zero: The temperature at which all translational motion of atoms and molecules ceases.

Absorption spectrum: An electromagnetic spectrum that shows dark lines which result from the passage of the electromagnetic radiation through an absorbing medium, such as the gases found in a star's atmosphere. The resulting *absorption lines* are characteristic of certain chemical elements and reveal much about the composition of the star's atmosphere. *See also* Electromagnetic spectrum, Emission spectrum, Spectrum.

Acceleration: The change in the velocity of an object divided by the time required for the change to occur; commonly expressed in units of meters per second squared.

Acceleration of gravity: The average acceleration of an object which is in free-fall near the Earth's surface, ignoring air resistance; this approximately constant acceleration has a value of 9.8 meters (32.15 feet) per second squared.

Accretion: The accumulation of matter that can result eventually in the formation of a planet or smaller-sized body. *See also* Condensation.

Accretion disk: A disk of material that spirals in toward a black hole or other compact object.

Achondrite: A stony meteorite that contains mostly silicate minerals and a small amount of metal formed from the cooling of molten rock.

Acquisition: The detection and tracking of an object, signal, satellite, or probe to obtain data or control the path of a spacecraft. *See also* Star tracker.

Active galaxy: A galaxy that contains a compact, highly energetic nucleus.

Aerodynamics: The study of the behavior of solid bodies, such as an airplane, moving through gases, such as Earth's atmosphere.

Aerography: The branch of meteorology that collects atmospheric data for the production of weather charts.

Aeronautics: The study of aircraft and aerodynamic spacecraft and the flight of these human-made objects in the Earth's atmosphere.

Aeronomy: The study of the physics and chemistry of the atmospheres of Earth and other planets.

Aerospace: The space extending from Earth's surface outward and beyond the atmosphere; also refers to the engineering discipline that designs and builds craft capable of operating in both the atmosphere and vacuum of space.

Airglow: A faint glow emitted by Earth which results from interaction between solar radiation and gases in the ionosphere, perceived from space as a halo around the planet. Airglow is known for interfering with Earth-based astronomical observations, making space-based telescopes desirable.

Albedo: The fraction of incident light that is reflected from planets, moons, and asteroids.

Alpha particle: A helium nucleus emitted during the radioactive decay of uranium, thorium, or other unstable nuclei.

Altitude: The distance of an object directly above a surface. Also, the arc or angular distance of a celestial object above or below the horizon.

Andesite: A type of volcanic igneous rock intermediate in composition and density between granite and basalt.

Anemometer: An instrument for measuring wind speed.

Angstrom (Å): One ten-thousand-millionth of a meter; a unit used to measure electromagnetic wavelengths.

Angular momentum: A property of a rotating body or a system of rotating bodies. For a discrete body, it is calculated as the product of the mass and the square of distance of the body from the axis about which rotation takes place. For a body with continuous distribution of mass, this calculation involves an integration over the body of distance squared times differential mass. In an analogy to translational motion where force results in a change in linear momentum, in rotational motion a torque results in a change in angular momentum. However, if a system has no net torque, then conservation of angular momentum applies, as it does for bodies orbiting the Sun or the planets. Angular momentum can also be expressed as the product of *moment of inertia*, a body's inertial resistance to rotational acceleration under the application of a torque, and *angular velocity*. When a planet is closest to the Sun, a point in its orbit when its distance is smallest, then its angular velocity is greatest. When a planet is farthest from the Sun, a point in its orbit when its distance is greatest, then its angular velocity is smallest. The latter two sentences describe Kepler's second law of planetary motion: The radius vector to a planet sweeps out equal areas in equal times.

Annular solar eclipse: An eclipse occurring when the apparent size of the Moon is slightly less than that of the Sun, so that even at points where the Moon is seen to move directly across the center of the Sun, the eclipse is not total, since a ring or annulus of the Sun can always be seen.

Anorthosite: A low-density igneous rock, consisting mostly of plagioclase feldspar, that comprises most of the outer crust of the Moon.

Anthropic principle: The philosophical viewpoint that the universe is structured so that galactic and stellar evolution inevitably will lead to the evolution of life, including intelligent life.

Antimatter: Matter composed of antiparticles: positrons (like electrons but with a positive charge), antiprotons (like protons but with a negative charge), antineutrons, and so forth. Small amounts of anti-hydrogen have been created for very brief times. The result of an encounter between a matter particle and its antimatter equivalent is mutual annihilation, with their mass converted into a burst of energy.

Antiparticles: The counterpart of elementary particles, having the same mass and spin as their corresponding particle, but with opposite electric charge and magnetic moment. *See also* Elementary particles.

Aphelion: The point in its orbit around the Sun at which an object, traveling an elliptical path, is farthest from the Sun.

Apoapsis: The point in one object's orbit around another at which the orbiting object is farthest away from the object being orbited.

Apocynthion: The point in an object's orbit around the Moon at which it is farthest away from the Moon.

Apogee: The point in an object's orbit around Earth at which it is farthest away from Earth.

Apolune: Apocynthion.

Apparent magnitude: The brightness of a star or other celestial body as seen from a point, such as Earth. The brightness is apparent because stars vary in their distance from Earth. *See also* Absolute magnitude, Luminosity.

Apparent motion: The path of movement of a body relative to a fixed point of observation.

Apparent size: The perceived or angular size of an object as viewed from a specific perspective, regardless of its true linear size.

Archean eon: The older of a two-part division of the Precambrian, the earliest era of geologic history. Also known as the Archeozoic.

Artificial satellite: A human-made satellite or object sent into orbit around a celestial body.

Asteroid: A small, solid body (also known as a minor planet) orbiting the Sun and ranging in size from about 1,000 kilometers in diameter down to a minimum of between 10 and 100 meters. The solar system contains many thousands of these objects, most of them having orbits in between the orbits of Mars and Jupiter, a region known as the asteroid belt. The Trojan asteroids are located 60° ahead of and behind Jupiter in its orbit. Some asteroids cross inside the orbit of the Earth.

Asteroid belt: The region between the orbits of Mars and Jupiter containing the majority of asteroids.

Asthenosphere: A region of Earth's upper mantle that has less rigid and probably plastic rock material that is near to but below its melting temperature.

Astrobleme: The remnant of a large impact crater on Earth; erosion will have altered the superficial appearance, but confirmation can be made from deeper structural damage and the presence of characteristically shattered and shocked rock.

Astronautics: The science and technology of spaceflight, including all aspects of aerodynamics, ballistics, celestial mechanics, physics, and other disciplines as they affect or relate to spaceflight. *See also* Aerodynamics, Aeronautics, Celestial mechanics.

Astronomical unit (AU): The mean distance between the centers of Earth and the Sun: 149,597,870 kilometers (92,955,630 miles). Used for measuring distance within the solar system.

Astronomy: The study of all celestial bodies and phenomena within the universe.

Astrophysics: The branch of astronomy dealing with the chemical and physical properties and behaviors of celestial matter and their interactions.

Ataxites: Iron meteorites that contain a very high nickel content.

Atmosphere: Any gaseous envelope surrounding a planet or star. Earth's atmosphere consists of five layers: the troposphere, stratosphere, mesosphere, thermosphere (which roughly coincides with the ionosphere), and exosphere.

Atmospheric pressure: The pressure exerted by a planet's atmosphere, decreasing with height and increasing with depth. Earth's atmospheric pressure at sea level is approximately 14.7 pounds per square inch, or 101,325 newtons per square meter, or 1.01325 bars. This pressure often is referred to as "one atmosphere" and can be used to describe the atmospheric pressure of other planets; for example, the atmospheric pressure at the surface of Venus is about 90 atmospheres, meaning 90 times Earth's atmospheric pressure at sea level.

Atom: The smallest particle of an element that can exist alone. Atoms consist of electrons (negatively charged particles) orbiting a nucleus made of protons (positively charged particles) and neutrons (particles with no net charge). The combinations of these particles determine the identity of the atom as a particular chemical element or isotope of that element. The number of protons (same as the number of electrons in a neutral atom) is the atomic number and therefore determines the chemical element as found in the periodic table; the total number of protons and neutrons in the nucleus is the atomic mass number and therefore identifies the particular *isotope* of that element.

Atomic number: The number of protons, or positively charged nuclear particles, in the nucleus of an atom.

Attitude: The orientation of a spacecraft or other body in space relative to a point of reference.

AU. *See* Astronomical unit.

Aurora: The colored lights appearing in the sky near the poles when charged particles issuing from the Sun become trapped in Earth's magnetic field. The arching, spiraling glows result from these particles interacting with atmospheric gases as they follow Earth's magnetic force lines.

Aurora australis: The aurora occurring near Earth's South Pole.

Aurora borealis: The aurora occurring near Earth's North Pole.

Avionics: The electronic devices used on board a spacecraft, or the development, production, or study of those devices.

Axis: The imaginary line around which a celestial body or human-made satellite rotates.

Axis tilt: A tilt in the pole-to-pole line about which a planet rotates, relative to the plane of the ecliptic. For example, Earth's axis tilt is 23.5°.

Azimuth: The arc, or angular distance, measured horizontally and moving clockwise, between a fixed point (usually true north) and a celestial object. *See also* Altitude.

Ballistics: The study of the motion of projectiles in flight, including their trajectories—especially important in the launching and course-correction maneuvers of spacecraft.

Band. *See* Frequency, Hertz.

Bar: A unit of pressure; one bar is defined as 100,000 Newtons per square meter, which is nearly the pressure of the Earth's atmosphere at sea level. *See also* Atmospheric pressure.

Barycenter: The center of mass of a system of two or more bodies.

Baryons: A class of particles that includes protons, neutrons, and the unstable hyperons; all baryons are composed of a combination of three quarks.

Basalt: A fine-grained, dark igneous rock composed chiefly of pyroxenes and feldspars, typically found at or near the surface of differentiated planets and moons.

Basin: A large, typically flat-bottomed crater formed by the impact of a very large body.

Beta decay: A radioactive decay process in which an electron and a neutrino are emitted from an atomic nucleus, transforming a neutron into a proton.

Big bang theory: The cosmological theory that the universe originated from a primordial ultra-hot, ultra-dense state, about thirteen to fifteen billion years ago. The big bang created space and time, matter and energy. Space rapidly expanded, and as it did so, energy and matter rapidly cooled, leading to the expanding universe we observe today. (The name "big bang" was first employed by an early opponent of this theory, Sir Fred Hoyle, who used it in a derogatory sense—but the name stuck.) *See also* Steady state theory.

Big crunch: The eventual recompression of all matter in the universe that may occur if the universe is closed.

Binary star: A star system composed of two stars orbiting their combined center of mass. Binary stars are termed "visual" if both components can be seen with a telescope, "spectroscopic" if their spectral lines are Doppler shifted alternately toward shorter and longer wavelengths by the orbital motion of the stars toward and away from us, and "eclipsing" if one star passes directly between us and the second star blocking at least some of its light.

Biosatellite: An artificial satellite carrying life-forms for the purpose of discovering their reaction to conditions imposed in the space environment.

Biotelemetry: The remote measurement and monitoring of the life functions (such as heart rate) of living beings in space, and the transmission of such data to the monitoring location, such as Earth.

Black dwarf star: A star that has cooled to the point that it no longer emits visible radiation; the end state of a white dwarf star. *See also* White dwarf star.

Black hole: A celestial body, predicted by Albert Einstein's general theory of relativity, in which gravity is so strong that nothing, not even light, can escape from it. Two types of black holes have been detected observationally. Stellar-mass black holes (or simply stellar black holes) have masses at least several times the Sun's mass and form from massive stars that collapse at the end of their energy-producing lives. Supermassive black holes have masses millions to billions of times the Sun's mass and occur at the centers of many galaxies, including our Milky Way. A possible third type of black hole, mini black holes, having about the mass of a mountain or less, may have formed in the very early universe in the aftermath of the big bang; mini black holes have not yet been detected.

Blueshift: An apparent shortening of electromagnetic wavelengths emitted from a star or other celestial object, indicating movement toward the observer. *See also* Doppler effect.

Bolometric magnitude: The brightness of a star as detected from above Earth's atmosphere and recorded in all wavelengths.

Bow shock: A shock wave (analogous to the shock wave preceding a supersonic aircraft) that is formed at the point where the solar wind (the stream of ionized gases flowing outward from the Sun) encounters a planet's magnetosphere.

Breccia: Rock composed of a random mixture of angular, broken fragments of other rocks and minerals.

Brown dwarf star: An intermediate-mass object, between a planet and a star, about ten to eighty times the mass of Jupiter, which formed like a star, but without enough mass to sustain nuclear fusion as a source of energy.

Caldera: A very large crater formed by the collapse of the central part of a volcano.

Caloris Basin: The largest known structure on Mercury; it is similar to the Moon's Imbrium Basin and was formed by a large impact.

Canopus: The brightest star in the sky after Sirius, visible south of 37° latitude. Canopus is often the target of a spacecraft's star tracker, which uses it as a reference point in steering a course toward the spacecraft's destination.

Carbon dioxide: A molecule consisting of one carbon atom and two oxygen atoms; it is a common constituent of planetary atmospheres.

Carbonaceous asteroid: An asteroid made up principally of carbon-based materials.

Carbonaceous chondrites: A class of stony meteorites found to contain large amounts of carbon in conjunction with other elements; used to date the solar system and to provide clues to the chemical composition of the early solar nebula.

Cassegrain telescope: A type of reflecting telescope, named for its inventor, Guillaume Cassegrain (1672). The telescope contains two mirrors: a concave mirror near its base, which reflects light from the sky onto a convex mirror above it; the convex mirror, in turn, reflects the light back down through a hole in the middle of the convave mirror to the focal point. The Hubble Space Telescope is a variant of the Cassegrain design called a Ritchy-Cretian system that uses aspheric mirrors.

Cassini division: The largest gap in Saturn's rings, directly viewable from Earth-based telescopes.

Catastrophism: The theory that the large-scale features of Earth were created suddenly by catastrophes in the past; the opposite of uniformitarianism. *See also* Uniformitarianism.

Celestial equator: A great circle on the celestial sphere 90° from the celestial poles, separating the northern and southern hemispheres of the sky.

Celestial mechanics: The branch of physics concerned with those laws which govern the motion (especially the orbits) of celestial bodies, both artificial and natural.

Celestial poles: Imaginary points around which the celestial sphere appears to rotate.

Celestial sphere: An imaginary sphere surrounding an observer at a fixed point in space (the sphere's center), with a radius extending to infinity, a *celestial equator* (a "belt" cutting the sphere into two even halves), and *celestial poles* (north and south). By reference to these points on the celestial sphere, the observer can describe the position of an object in space.

Celsius scale: A temperature scale, named for its inventor, Anders Celsius (1701-1744), which sets the freezing point of water at 0° and the boiling point at 100°. Also referred to as the centigrade scale, its increments correspond directly to kelvins. To convert kelvins to degrees Celsius, subtract 273.15. *See also* Kelvin.

Centrifugal force: The pseudo-force tending to impel a body outward from a center of rotation, equal and opposite to the *centripetal force* caused by the inertia of the body.

Centrifuge: A device for whirling objects or human beings at high speeds around a vertical axis, exerting centrifugal force to test spacecraft hardware or train astronauts to withstand the forces of launch.

Cepheid variable: A massive star that has passed its main sequence phase (the greater part of its lifetime) and has entered a transitional phase in its evolution, during which the star expands and contracts, pulsating in brightness. By measuring the star's period of pulsation and estimating from that its absolute magnitude, then comparing the absolute magnitude to the apparent magnitude, astronomers find the distance of the star.

Chandrasekhar limit: The maximum possible mass for a white dwarf star, calculated by Subramanyan Chandrasekhar in 1931 as approximately 1.4 solar masses (later modified upward for rapidly rotating white dwarf stars). When the star's mass exceeds the Chandrasekhar limit, gravity compresses it into a neutron star.

Charge: A property of matter defined by the excess or deficiency of electrons in comparison to protons. Negative charge results from excess electrons; positive charge, from a deficiency of electrons.

Chemical evolution: The synthesis of amino acids and other complex organic molecules—the precursors of living systems—by the action of atmospheric lightning, solar ultraviolet radiation on atmospheric gases, and other sources of energy.

Cherenkov light: Light emitted by a particle that is exceeding the speed of light in the medium through which it is traveling.

Chert: A hard rock of minutely crystalline, and often partly hydrous, silica.

Chondrite: A stony meteorite containing glassy spherical inclusions called chondrules, which are usually composed of iron, aluminum, or magnesium silicates.

Chromosphere: The layer of the solar atmosphere between the photosphere and the corona, several thousands of kilometers thick, which is visible only when the photosphere is obscured, as during a solar eclipse or by using special filters. The term also applies to corresponding regions of other stars.

Circle of illumination: The circle on Earth's surface that bisects Earth and separates the sunlit half from the shadowed half.

Circular orbit: An orbit described by an orbiting body that maintains a constant distance about the body around which it orbits. The orbit's eccentricity is zero as the orbit's pair of foci are identical.

Climate: The sum total of the prevailing long-term weather conditions of an area, determined by such factors as latitude, altitude, and location.

Closed universe: A universe in which all matter will eventually recompress into a tiny volume of space.

Coesite: A high-density type of quartz formed under the pressures and temperatures involved in impact cratering.

Coma: The gaseous envelope surrounding the head of the comet and consisting of evaporated gases from the comet's nucleus.

Comet: A luminous celestial object orbiting the Sun, consisting of a nucleus of water ice and other ices mixed with solid matter, and, as the comet approaches the Sun, a growing coma and tail. The coma is a collection of gases and dust particles that evaporate from the nucleus and form a glowing ball around it; the tail forms as these materials are swept away from the nucleus. Comets appear periodically, depending on the parameters of their solar orbits.

Comet period: The time required for a comet orbiting the Sun to complete a single orbit.

Condensation: A condition in the early solar nebula when hot gases cooled to form solids; a preliminary stage to accretion. *See also* Accretion.

Condensation temperature: The temperature at which gases of the primitive solar nebula condensed into solid particles.

Conjunction: The alignment of two planets or other celestial bodies so that their longitudes on the celestial sphere are the same. *Inferior conjunction* occurs when Mercury or Venus passes between the Earth and the Sun; *superior conjunction* occurs when Mercury or Venus passes behind the Sun as seen from Earth.

Conservation law: A rule of physics that states that the total value of some quantity does not change.

Conservation of angular momentum: The principle that the total angular momentum of a body or group of objects in a system remains constant in the absence of any external torque.

Constellation: A collection of stars which form a pattern as seen from Earth. The stars in these groupings are often quite distant from one another, their main common characteristic being the illusory picture they form (such as the Big Dipper, or Ursa Major) against the backdrop of the night sky. Constellations provide points of reference for astronomers and other stargazers.

Convection: A flow of material resulting from temperature differences that cause warm, light material to rise and cool, dense material to sink.

Convective equilibrium: A state of stellar stability characterized by a fluid, convective transfer of energy from hotter inner regions to cooler outer layers.

Coorbital satellites: Bodies that share the same orbit; these bodies are in a 1:1 orbital resonance.

Core: The central portion of any celestial body, such as planets and stars. Earth's core is located 2,900 kilometers below the planet's surface.

Core-diffracted phases: Those elastic waves incident at the Earth's outer core at a grazing angle that are diffracted and arrive within the shadow zones for direct waves.

Core-reflected phases: Elastic waves that are reflected from the Earth's core-mantle boundary.

Core sample: A sample of rock and soil taken from Earth, the Moon, or another terrestrial planet by pressing a hollow cylinder down into the planet's surface.

Core-transmitted phases: Elastic waves that travel through the Earth's core.

Coriolis force: A pseudo-force that deflects moving objects to the right (Northern Hemisphere) or to the left (Southern Hemisphere) because of the Earth's rotation. The term describes an inertial movement in a noninertial reference frame.

Corona: The outermost portion of the Sun's atmosphere, extending like a halo outward from the Sun's photosphere. The corona consists of extremely hot ionized gases that eventually escape as solar wind. The term is also used to refer to the corresponding region of any star's atmosphere.

Coronagraph: A device for viewing the solar (or another star's) corona, consisting of a solar telescope outfitted with an occulting mechanism to obscure the photosphere, as during a solar eclipse, so that the corona is more easily perceived.

Cosmic background radiation: Relic radiation produced a few hundred thousand years after the big bang, when the early expanding universe had cooled sufficiently (down to a few thousand kelvins) for electrons to join with protons to form neutral hydrogen atoms. Before that time, the free electrons made the early universe opaque to electromagnetic radiation. Once the free electrons were incorporated into neutral hydrogen atoms, the universe became transparent to electromagnetic radiation, and photons were free to fly throughout the expanding universe. When the universe first became transparent, the photons had a distribution of energies and wavelengths characteristic of a thermal, blackbody source at a temperature of several thousand kelvins. As the universe expanded, the photon energies decreased and the wavelengths increased. Today the radiation is observed greatly redshifted (by about a factor of 1,000) into the microwave portion of the spectrum, with wavelengths of approximately 1 million nanometers or 1 millimeter, appearing like thermal, blackbody radiation from a source with a temperature of approximately 3 kelvins.

Cosmic dust: Tiny solid particles found throughout the universe, thought to have originated from the primordial universe, the disintegration of comets, the condensation of stellar gases, and other sources. Also known as interstellar dust.

Cosmic microwave background radiation: Microwave radiation from the glowing of the hot early universe, now cooled by the expansion of the universe to a temperature of about 2.7 kelvins.

Cosmic radiation: Atomic particles, also known as cosmic rays, that are the most energetic known, consisting mainly of protons, along with electrons, positrons, neutrinos, gamma-ray photons, and various atomic nuclei. These particles emanate from a number of sources, both within and beyond the Milky Way, and they bombard atoms in Earth's atmosphere to produce showers of secondary particles such as pions, muons, electrons, and nucleons. If a primary cosmic-ray particle is sufficiently energetic when it hits an atmospheric atom, the secondary particles can reach Earth's surface.

Cosmic string: A hypothetical early concentration of energy that initiated formation of galaxy filaments.

Cosmogony: The study of the origin and nature of the solar system.

Cosmology: The study of the large-scale structure of the universe and its movements, origin, evolution, and ultimate fate.

Coulomb repulsion: A repulsive electrical force experienced between charged particles of similar sign, such as nuclei of elements as a result of their similar positive charges, and that must be overcome for fusion to take place. The magnitude of this force is directly proportional to the product of the charges of the two interacting objects and inversely proportional to the distance between the centers of these two charged bodies or particles.

Covariant: Interdependent according to a particular mathematical rule, such as that which connects the space coordinates and time of an event in relativity.

Crab nebula: An important supernova remnant that contains a pulsar at its center; also known as M1.

Crater: A depression in the surface of a planet or moon caused by the force of a meteorite's fall. Also, the depression that forms at the mouth of a volcano.

Crater morphology: The structure or form of craters and the related processes that produced them.

Crust: The outermost layer of a planet, generally composed of materials removed from the interior by chemical and physical processes.

Crustal differentiation: The process resulting in the origin of continental and oceanic crust through remelting of original, heavier crust.

Crystal: A solid made up of a regular periodic arrangement of atoms or molecules; its form and physical properties express the repeat units of the structure.

Crystallization: The process by which minerals are formed at various temperatures and pressures, resulting in an orderly arrangement of their atoms.

Cumulate: An igneous rock composed chiefly of crystals which accumulated by sinking or floating from a magma.

Cumulus crystals: Dense minerals within liquid magma that accumulate by gravity upon the floor of a magma chamber.

Curie temperature: The temperature above which a permanently magnetized material loses its magnetization.

Curved space: A space which does not obey the rules of Euclidean geometry; may be positively curved or negatively curved.

Cyclotron radiation: The radiation produced by charged particles as they spiral around magnetic field lines at extremely high speeds.

Dark cloud: An interstellar cloud of dust with sufficient density to block the passage of starlight.

Dark energy: The name given to the hypothetical energy that is causing the expansion of the universe to accelerate. According to modern cosmological models, it should amount to approximately 70 percent of all the energy and matter in the universe.

Dark matter: Mass in the universe that does not give off any form of electromagnetic radiation and is thus invisible but is known by its gravitational influence.

Day: The interval of time between two successive passages of the Sun over the meridian of a planet.

Declination: The angular distance north or south from the celestial equator measured along a circle passing through the celestial poles.

Deep space: Regions of space beyond the Earth-Moon system.

Deep space probe: A device launched beyond the Earth-Moon system that is designed to investigate other parts of the solar system or beyond. Sometimes called an *interplanetary space probe* in reference to spacecraft investigating the planets and the space between them.

Degenerate electron gas: An assembly of electrons occupying all the quantum mechanically allowed space, occurring within white dwarf stars; the pressure in an electron degenerate gas does not depend on temperature but only on the density.

Degradation: Erosion from all processes, including wind, rain, and other mechanisms.

Density: The amount of matter that is contained within a given volume of space; mass per unit volume (grams per cubic centimeter or kilograms per cubic meter).

Differentiation: A process that separates materials by some criteria, such as the heating and melting of a terrestrial protoplanet in which denser material sinks toward the center to become the core and less dense material rises toward the surface to become the crust.

Digital imaging. *See* Imaging.

Dipole field: Electrically, the field shape produced prototypically by a pair of electrically charged particles of equal magnitude but opposite sign separated by a small distance relative to the distance at which the dipole is being observed. Magnetically, the field created by the north and south poles which cannot be found alone, but must always occur in pairs.

Dirty snowball model: The model of a comet's nucleus as a dirty snowball, consisting of various frozen ices containing sand and dust grains.

Discontinuity: An abrupt change in some property at a boundary, such as the change in speed of seismic waves at the boundary between the Earth's crust and mantle (called the Mohorovičić discontinuity).

Dissociation: The breaking up of a compound into simpler components, such as the separation of a molecule into its constituent atoms.

Diurnal: Occurring daily. The diurnal motion of a planet or other celestial body is its daily path across the sky as seen from a fixed point such as Earth.

Doppler effect: The effect of an object's motion toward or away from an observer on the wavelength and frequency of electromagnetic or sound waves that it emits. If the object is moving away from the observer, the observed wavelength is longer and the observed frequency is lower compared to what the object emitted. If the object is moving toward the observer, the observed wavelength is shorter and the observed frequency is higher compared to what the object emitted. The faster the speed of the object, the more the wavelength and frequency are shifted away from their emitted values. Since red light is at the long wavelength end of the visible spectrum, and blue and violet light are at the short wavelength end of the visible spectrum, an object moving away from the observer will appear redder (called a redshift) and an object moving toward the observer will appear bluer (called a blueshift). A source of sound will be heard at a higher pitch as it approaches and at a lower pitch as it moves away.

Double star. *See* Binary star.

Draconic month: The period from one crossing of the ecliptic plane by the Moon to the next (in the same direction), about 27.21 days.

Dynamo effect: The movement of an electrical conductor through a magnetic field, producing an electrical current that in turn generates a magnetic field.

Earth day: Twenty-four hours, or the time required for Earth to complete one rotation on its axis.

Earth-emitted radiation: The portion of the electromagnetic spectrum, from about 4 to 80 microns, in which Earth emits about 99 percent of its radiation.

Earth-orbital probe: An uncrewed spacecraft carrying instruments for obtaining information about the near-Earth environment.

Earth tide: The slight deformation of Earth resulting from the same forces that cause ocean tides, those that are exerted by the Moon and the Sun.

Eccentricity: The degree to which an ellipse (or orbital path) departs from circularity. Eccentricity is characterized as "high" when the ellipse is very elongated. For conic sections, eccentricity is zero for a circle, is between zero and one for an ellipse, is unity for a parabola, and is greater than one for a hyperbola.

Eclipse: The obscuring of one celestial body by another. In a lunar eclipse, Earth's shadow obscures the Moon when Earth is situated directly between the Sun and Moon. In a solar eclipse, the Moon is situated between the Sun and Earth in such a way that part or all of the Sun's light is blocked; the total blockage of sunlight (with the exception of the Sun's corona) is called a total eclipse of the Sun.

Ecliptic: The apparent annual path of the Sun on the celestial sphere.

Ecliptic plane: The plane in which Earth orbits the Sun. From Earth, the ecliptic plane is perceived as the Sun's yearly path through the sky.

Ecology: The relationship between organisms and their environment.

Effective temperature: The temperature of a blackbody (an ideal thermal radiator) that emits the same total flux of electromagnetic radiation as the object under consideration.

Einstein cross: A collection of four images of a distant object formed by a massive object acting as a gravitational lens.

Einstein ring: The images of a distant extended object (such as a galaxy) elongated into arcs when viewed past a mass that bend the light rays.

Ejecta: Material thrown out of a volcano during eruption. Also, the material ejected from the crater made by a meteoric impact.

Electrolysis: A process whereby water is broken down into oxygen and hydrogen.

Electromagnetic radiation: A phenomenon processing both wave and particle characteristics (referred to as wave-particle duality). In some situations, it behaves as waves of electric and magnetic fields oscillating perpendicular to each other and the direction of propagation, and traveling at the speed of light. In other situations it behaves as particles called photons that carry energy and momentum, but have no rest mass. The energy and momentum of a photon are directly proportional to the frequency and inversely proportional to the wavelength of the electric and magnetic field oscillations. *See also* Electromagnetic spectrum.

Electromagnetic spectrum: The continuum of all possible electromagnetic wavelengths, from the longest, radio waves (longer than 0.3 meter), to the shortest, gamma rays (shorter than 0.01 nanometer). The shorter the wavelength, the higher the frequency and the greater the energy. Within the electromagnetic spectrum is a range of wavelengths that can be detected by the human eye, visible light. Its wavelengths correspond to colors: Red light emits the longest-wavelength visible radiation; violet light, the shortest-wavelength radiation. None of these types of electromagnetic radiation is discrete; each blends into the surrounding forms. Detection of nonvisible radiation by special instruments (used in such branches of astronomy as infrared astronomy and X-ray astronomy) reveals much about the behavior of celestial bodies and the origins of the universe.

Electron: An elementary particle which carries a negative charge and a mass about one eighteen-thousandth of a proton. The number of electrons in an atom in its neutral state is determined by its atomic number and therefore is the same as the number of protons in its nucleus. If an atom accepts or loses one or more electrons then it becomes ionized and has a net negative or positive charge, respectively.

Electron volt: The amount of energy gained by an electron when it is accelerated through an electrical potential difference of one volt; one electron volt (symbol eV) equals 1.6×10^{-19} joules.

Electrophoresis: A process for separating cells using a weak electric field, more easily accomplished in space than on Earth.

Element: The simplest chemical substance, made up of atoms of identical atomic number; every element has a unique atomic number.

Elementary particles: The smallest units of matter, characterized by electrical charge, mass, and angular momentum. Among elementary particles are electrons, neutrons, protons, neutrinos, the various mesons, and their corresponding antiparticles (which form antimatter). Photons, the smallest units of electromagnetic radiation, are also considered as elementary particles.

Ellipse: An oval-shaped geometric figure traced by a point moving so that the sum of its distances from two other points (called the foci) remains constant. The eccentricity of an ellipse varies between zero and one. A circle is a special case of an ellipse in which the two foci are the same point, i.e. the center of the circle.

Elliptical orbit: An orbit which departs from circularity, as most orbits do. A *highly elliptical orbit* is one whose apoapsis is much greater than its periapsis, resulting in an orbit that traces out an elongated ellipse.

Emission nebula: A cloud of gaseous material hot enough to be observed by its own emitted light.

Emission spectrum: Also known as emission line spectrum or bright line spectrum, a spectrum consisting of a series of bright, colored lines at certain specific wavelengths, emitted by a diffuse glowing gas. Each chemical element in gaseous form emits its own characteristic set of bright, colored lines at a particular set of wavelengths.

Enantiomer: A particular version of the same kind of asymmetric chemical compound, such as sugars and amino acids, that may be left- or right-handed and so polarize light in a clockwise (D enantiomer) or counterclockwise (L enantiomer) direction, respectively.

Energy: The ability to do work.

Enstatite chondrites: A rare group of meteorites composed of recrystallized agglomerates whose textural features and mineralogy represent conditions of thermal metamorphism under reducing conditions.

Eolian erosion: A mechanism of erosion or crater degradation caused by wind.

Ephemerides: Calculations showing predicted positions of celestial objects, which can include the Sun, Moon, and planets, and times of eclipses, sunrise, and sunset; ephemerides may also plot a particular object, such as a comet or minor planet.

Epicenter: The region at a planet's surface directly above the focus, or hypocenter, of a seismic quake.

Equatorial orbit: An orbit that directly overlies a planet's equator.

Erg: The amount of kinetic energy of a mass of 2 grams moving at 1 centimeter per second; a mosquito in flight possesses about one erg of energy.

Escape velocity: The speed at which an object must travel to escape the gravitational attraction of a celestial body. In order for a spacecraft to leave Earth orbit, for example, its engines must exert enough in-orbit thrust to achieve escape velocity.

Event: A fundamental "point" of space-time, specified not only by a place but also by a time of occurrence.

Event horizon: The boundary beyond which an observer cannot see. Also, the boundary beyond which nothing can escape from a black hole, where escape velocity equals the speed of light and thus nothing, not even light, can escape. Therefore, the event horizon is theoretically the spherical delineation of a black hole. *See also* Escape velocity.

Exclusion principle. *See* Pauli exclusion principle.

Exobiology: The study of the conditions for and potential existence of life-forms beyond Earth.

Exosphere: The outermost region of the atmosphere.

Extraterrestrial life. *See* Exobiology.

F region. *See* Thermosphere.

Faculae (*sing.* facula): Bright spots or streaks on the solar photosphere associated with the magnetic field of the Sun.

False-color image: An image resembling a photograph, created from data collected by instruments (such as an infrared sensor) aboard a spacecraft and deliberately assigned unnatural colors in order to make nonvisible radiation visible or to highlight distinctions. *See also* Imaging.

Faraday cup: A probe flown on spacecraft to determine the energy (and therefore the velocity) of the plasma in the vicinity of the probe.

Fireball: A very large and bright meteor that often explodes with fragments falling to the ground as meteorites; sometimes called a bolide.

Fission (atomic): The splitting of an atomic nucleus into less massive parts, resulting in a great release of energy.

Fission tracks: Regions of damage to a crystal along the path taken by a moving ion, usually a fragment resulting from fission decay or a cosmic ray.

Flat space: A space of any number of dimensions which obeys the rules of Euclidean geometry.

Flight path: The trajectory of an airborne or spaceborne object relative to a fixed point such as Earth.

Fluid mechanics: The study of the behavior of fluids (gases and liquids) under various conditions, including that of microgravity in spaceflight. Understanding fluid mechanics in space is important to the technology of spaceflight and may have applications on Earth as well.

Fluorescence: Re-emission of visible light due to absorption of electromagnetic radiation of a higher energy (usually ultraviolet light) from an external source.

Flyby: A close approach to a planet or other celestial object, usually made by a probe for the purpose of gathering data; the maneuver does not include orbit or landing. Also used to refer to a mission which undertakes a flyby.

Focal ratio (f-number): The ratio of (1) the distance between the center of a lens or mirror and its point of focus (focal-length) and (2) the aperture, or diameter, of the lens.

Focus (*pl.* foci): (1) The region within Earth from which earthquake waves emanate; also called its hypocenter. (2) The two points that help define an ellipse or hyperbola geometrically. (3) The point at which the light refracted by a lens or reflected by a curved mirror forms an image.

Force: A physical phenomenon capable of changing the momentum of an object; the four *fundamental forces* in the universe are gravity, electromagnetism, the strong nuclear force, and the weak nuclear force.

Frame of reference: A particular position, moving or stationary, from which objects and events are observed.

Fraunhofer lines: Prominent absorption lines in the Sun's spectrum, first observed by Joseph von Fraunhofer in 1814, indicating the presence of certain elements in the Sun's atmosphere. Also used to refer to such absorption lines in other stars' spectra.

Free return trajectory: An orbital flight path which allows a disabled spacecraft to reenter Earth's atmosphere without assistance.

Frequency: The number of times an event recurs within a specific period of time. Frequency characterizes all wave phenomena (such as sound, seismic, and electromagnetic waves) and is determined by dividing the wave speed by its wavelength. Frequency is measured in hertz or multiples of hertz.

Fusion: A thermonuclear reaction in which the nuclei of light elements are joined to form heavier atomic nuclei, releasing energy. It is the process whereby stars form the elements with atomic numbers up to that of iron. It is also the process whereby the stars formed the elements with atomic numbers up to iron.

Gabbro: An igneous rock consisting mostly of pyroxene, feldspar, and often olivine, and containing less than 55 percent silica.

Gain: The increase in power of a transmitted signal as it is picked up by an antenna.

Galactic cluster: An archaic, obsolete, ambiguous designation that could refer to a cluster of several hundred to several thousand stars in the main disk of a spiral galaxy, or to a cluster of galaxies. The star clusters now are referred to as *open clusters* or *open star clusters* to distinguish them from globular star clusters. The clusters of galaxies now are referred to as *galaxy clusters*. In both cases, the clusters are held together by the mutual gravitational attraction of their members.

Galaxy: A collection of millions to trillions of stars, gas, and dust gravitationally bound together.

Gamma radiation: The most energetic form of electromagnetic radiation, with wavelengths less than 0.01 nanometer. The ability of gamma rays to penetrate the interstellar matter of the universe makes them especially valuable to astronomers.

Gamma-ray astronomy: The branch of astronomy that investigates gamma radiation and its sources. Gamma-ray observatories sent into orbit have included the Orbiting Solar Observatory, SAS 2, COS-B, the Compton Gamma Ray Observeratory, and Fermi.

Gaps: Spaces in planetary rings caused by gravitational interactions between the planet, its moons, and the ring particles.

Gegenschein: A patch of faint light about 20°

across, visible from the nightside of Earth at a point opposite the Sun in the ecliptic plane, and caused by the back-scattering of sunlight from interplanetary dust grains in the ecliptic plane. *See also* Zodiacal light.

Geiger counter: A device that detects high-energy radiation (in the form of particles and photons) by means of a tube containing gas and an electric current. The radiation causes the gas to ionize, which is transmitted to the current and detected as a sound or a needle jump.

General relativity: Albert Einstein's theory of gravitation, which extends Sir Isaac Newton's theory by stating that matter curves space and time; gravity is explained as matter moving along shortest paths, called geodesics, in the curved space-time.

Geocentric orbit: An orbit with Earth as the object orbited.

Geochemical sinks: The means by which elements and compounds are removed from the crust, atmosphere, and oceans to be recycled in active chemical cycles.

Geochronology: The study of the time scale of Earth; it attempts to develop methods that allow the scientist to reconstruct the past by dating events such as the formation of rocks.

Geodesy: The science concerned with the size and shape of Earth and its gravitational field.

Geodynamo theories: Theories that explain the cause of Earth's magnetic field and its secular variation in terms of electric charges in the Earth's molten metallic outer core.

Geomagnetic elements: Measurements that describe the direction and intensity of Earth's magnetic field.

Geomagnetism: The external magnetic field generated by forces within Earth; this force attracts magnetic materials, inducing them to line up their magnetic moments along the Earth's magnetic field lines.

Geometry: A set of rules which describes the structure of a region of space; traditional geometries described space as flat, while relativistic geometries describe it as curved.

Geomorphology: The study of landforms on planetary surfaces and the processes responsible for their origin.

Geostationary orbit: A type of geosynchronous orbit which is circular and lies in Earth's equatorial plane, at an altitude of approximately 36,000 kilometers (22,320 miles). As a result, a satellite in geostationary orbit appears to hover over a fixed point on Earth's surface. *See also* Geosynchronous orbit.

Geosynchronous orbit: A geocentric orbit with a period of 23 hours, 56 minutes, 4.1 seconds, equal to Earth's rotational period. Such an orbit is also geostationary if it lies in Earth's equatorial plane and is circular. If inclined to the equator, a geosynchronous orbit will appear to trace out a figure eight daily; the size of the figure eight will depend on the angle of inclination. These orbits are used for satellites whose purpose it is to gather data on a particular area of Earth's surface or to transmit signals from one point to another. Communications satellites are geosynchronous.

Geothermal: Pertaining to the heat of the interior of a planet.

Gigahertz. *See* Hertz.

Glass: A solid consisting of a disordered pattern of atoms, which represents a rapid cooling from a molten state; in meteorites, it is found in chondrules and within the matrix as fragments.

Globular clusters: Spherically shaped congregations of tens to hundreds of thousands of stars that occur throughout the universe, although more often near elliptical galaxies than spiral galaxies such as the Milky Way. It is believed that globular clusters contain the oldest stars.

Granite: A silica-rich igneous rock light in color, composed primarily of the mineral compounds quartz and potassium- and sodium-rich feldspars.

Gravitation *or* Gravity: The force of attraction which exists between two bodies, such as Earth and the Moon. In 1687, Sir Isaac Newton described this force as proportional to the product of the masses of the two bodies and as inversely proportional to the distance between them squared. Although gravitation is the weakest of the naturally occurring forces (the others being electromagnetic and nuclear in nature), it has the broadest range and is responsible for much celestial movement, including orbital dynamics.

Gravitational constant, *G*: The constant ($6.67\ 10^{-11}$ newton meters^2/kilograms^2) in Newton's law of gravity that determines the gravitational attraction between two bodies.

Gravitational contraction: The shrinking of an object, such as a protostar or protoplanet, due to the gravitational attraction of every bit of matter in the object for every other bit of matter in the object.

Gravitational differentiation: The separation of minerals, elements, or both as a result of the influence of a gravitational field wherein heavy phases sink or light phases rise through a melt.

Gravitational field: The acceleration field (force per unit mass) created by the mass of a celestial body. It assigns to every point in space around the body both an acceleration and a direction of motion that would be experienced by another body placed in the vicinity of the gravitating body that sets up this field.

Gravitational force: An attraction that acts on all masses, causing weight and orbital motion.

Gravitational lens: A large mass that bends the light from distant objects.

Gravitational mass: That property of an object that produces its gravitational attraction for other objects; it is the mass that appears in Newton's formula for the gravitational force between two masses $F = Gm_1m_2/r^2$ (force equals Newton's universal gravitation constant *G* times the product of the masses of the two gravitationally interacting bodies divided by the square of the distance between the centers of these two masses). In contrast, inertial mass is the property of an object that resists any change in its motion. In all tests, the gravitational mass and inertial mass of an object are individual.

Gravitational potential energy: The energy a body has by virtue of its location in a gravitational field; in a star, the extended outer atmosphere stores a large amount of gravitational potential energy, which can be released during contraction.

Gravity assist: A technique, first used with the Mariner 10 probe to Mercury, whereby a spacecraft uses the gravitational and orbital energy of a planet to gain energy to achieve a trajectory toward a second destination or to return to Earth.

Grazing incidence: Reflection where the incoming radiation strikes at an extremely shallow angle (typically less than 1°).

Great Red Spot: A vast, oval-shaped cloud system occurring at 22° south latitude in Jupiter's atmosphere, rotating counterclockwise. Its red color comes from an unknown substance that convection pulls to the surface; the substance absorbs violet and ultraviolet radiation and consequently delivers a red hue to the Spot. The Great Red Spot has been observed for more than three centuries.

Greenhouse effect: The heating of a planet's surface and lower atmosphere as a result of trapped infrared radiation. Such radiation becomes trapped when there is an excess of certain gases such as carbon dioxide in the atmosphere, which absorbs and reemits infrared radiation rather than allowing it to escape. As a result, the atmosphere acts like a greenhouse, heating the planet.

Groundwater: The water that occurs in the subsurface of a planet; it particularly applies to the subsurface zone that is saturated with such water.

Guest star: Another term, in ancient Chinese astronomical records, for a nova, or "new" star.

Gyroscope: A device which uses a rapidly spinning rotor to assist in stabilization and navigation.

H I region: A region in which the element hydrogen exists primarily in the form of neutral atoms.

H II region: A region in which the element hydrogen is ionized, existing as separate protons and electrons, necessarily at a higher temperature than an H I region.

Hadron: Any particle that participates in the strong interaction; hadrons are divided into baryons, which obey the Pauli exclusion principle, and mesons, which do not.

Heat budget: The balance between incoming short-wavelength solar radiation and outgoing long-wavelength infrared terrestrial or planetary radiation.

Heat death: The eventual loss of all usable energy in the universe that will occur if the universe is open (that is, if it will expand forever).

Heliocentric orbit: An orbit with the Sun at its center or at one of its two foci.

Heliopause: The border between the solar system and the surrounding interstellar space, where the solar wind gives way to interstellar matter and winds.

Heliosphere: The region of interplanetary space extending outward from the Sun in which the solar magnetic field controls the behavior of charged particles.

Hertz: An SI unit of frequency, equaling one cycle per second. Multiples include kilohertz (10^3 hertz), megahertz (10^6 hertz), and gigahertz (10^9 hertz).

Hertzsprung-Russell diagram: A graph widely used in astronomy that depicts stellar properties such as luminosity or absolute magnitude versus spectral type or effective photospheric temperature. Named after Ejnar Hertzsprung and Henry Norris Russell, who independently devised the graph in the early twentieth century.

Hexahedrites: Iron meteorites that contain less than 6 percent nickel content; they usually consist of large single crystals of kamacite and may show Neumann bands when polished.

High-Earth orbit: Any Earth orbit at a relatively great distance from Earth, such as the geosynchronous orbits of telecommunications satellites.

High-gain antenna: A single-axis, strongly directional antenna that is able to receive or transmit signals at great distances.

Highlands: Densely cratered regions on the lunar surface, which when seen with the naked eye take on a pale color; they are primarily anorthositic breccia.

Homogeneity problem: The difficulty of reconciling observations of the extreme uniformity of the cosmic microwave background radiation with the early inhomogeneity required to account for large-scale structures in the universe, such as galaxy clusters and voids.

Homogeneous accretion theory: One of two major theories on Earth's differentiation: that differentiation occurred after the Earth had formed through accretion of debris. *See also* Inhomogeneous accretion theory.

Horizon: The line formed where land meets sky, from the perspective of an observer. In astronomy, the horizon also means the circle on the celestial sphere that is formed by the intersection of the observer's horizontal plane with the sphere. The *particle horizon* is the theoretical limit beyond which particles cannot yet have traveled.

Horizontal branch: A stage of helium fusion in the cores of stars; in star clusters, this appears as a nearly horizontal grouping in the Hertzsprung-Russell diagram.

Hour circle: A great circle on the celestial sphere passing through the celestial poles.

Hubble constant: The constant of proportionality between the recessional velocity of galaxies and their distances in Hubble's Law.

Hubble's law: The principle, articulated in 1929 by Edwin Hubble, that the galaxies are moving away from one another at speeds proportional to their distance, that is, uniformly across time. Hubble deduced this principle from observations of the redshifts in galactic spectra. Along with the discovery of the cosmic microwave background radiation by Arno Penzias and Robert Wilson in 1965, Hubble's law forms the basis for the big bang theory of the expanding universe. *See also* Big bang theory, Doppler effect, Redshift.

Hydrocarbons: Molecules containing hydrogen, carbon, and oxygen.

Hydrostatic equilibrium: A state of stellar stability characterized by a balance between gravitational contraction and thermal pressure expansion.

Igneous rock: A rock formed by the cooling of molten material.

Imaging: The process of creating a likeness of an object by electronic means.

Impact basin: A large circular basin produced by a meteorite impact.

Impact breccia: Angular, fragmental rock produced by meteorite impact.

Impact crater: A generally circular depression formed on the surface of a planet by the impact of a high-velocity projectile such as a meteoroid, asteroid, or comet.

Inclination. *See* Orbital inclination.

Inclusions: In meteorites, rounded or irregular shapes that have textures and mineralogies suggestive of unmelted aggregates of solid particles, indicative of a primitive origin.

Index of refraction: A physical parameter of matter that describes the degree to which a ray of light will bend upon entering that medium. Specifically the index of refraction n is given by the speed of light in vacuum divided by the speed of light within the medium; since the latter is always slower than the former, the index of refraction of a medium is always greater than one, being equal to unity only for vacuum.

Inertia: The tendency of a body to stay at rest or in uniform motion unless acted upon by an unbalanced force.

Inertial mass: That property of an object that resists any change in its motion; it is the mass that appears in Newton's law of motion $F = ma$ (force equals mass times acceleration). In contrast, gravitational mass is the property of an object that makes it attract other masses. In all tests, the inertial mass and gravitational mass of an object are indistinguishable.

Inflation: The quantum cosmological scenario, which postulates that, almost immediately after the big bang, the universe underwent an enormous expansion.

Infrared: A component of the electromagnetic spectrum that is found just beyond the red part of the visible spectrum.

Infrared astronomy: The branch of astronomy which examines the infrared emissions of stars and other celestial phenomena. Studying the infrared emissions tells astronomers much about the composition and dynamics of their sources. Because infrared radiation cannot readily penetrate most of Earth's atmosphere, infrared astronomy is done at special observatories located at great altitude, such as at Mauna Kea in Hawaii or with orbiting infrared telescopes such as the Spitzer Space Telescope.

Infrared radiation: Electromagnetic radiation of wavelengths from 1 to 1,000 microns, wavelengths that occur beyond the red end of the visible portion of the electromagnetic spectrum.

Infrared spectrometer: A spectrometer that takes spectra of infrared radiation.

Inhomogeneous accretion theory: One of two major theories on Earth's differentiation: that differentiation occurred while the Earth was accreting debris because denser debris formed first. *See also* Homogeneous accretion theory.

Intercrater plain: Terrain consisting of gently rolling plains littered with small secondary craters; the crater density is higher than in a smooth plain.

Interferometer: An instrument that uses interference of electromagnetic waves to produce images with better resolution (sharper detail).

Interferometry: A data acquisition technique which uses more than one signal receiver (such as a series of radio telescopes). Signals are combined to form one highly detailed image. *See also* Interferometer, Very long baseline interferometry.

Intergalactic medium: Matter that exists between galaxies.

International Geophysical Year (IGY): The eighteen-month period from July, 1957, to December, 1958, during which many countries cooperated in the study of Earth and the Sun's effect on it. During this time, the space age can be said to have begun with the launch of Sputnik 1 on October 4, 1957.

Interplanetary space probe. *See* Deep space probe.

Interstellar dust. *See* Cosmic dust.

Interstellar medium: The material that lies between the stars; it consists mainly of grains of dust and gas, mostly hydrogen, along with heavier elements released by supernova explosions.

Interstellar wind. *See* Solar wind, Stellar wind.

Interval: A measure of the separation in spacetime between two events; intervals may be timelike, spacelike, or null (lightlike).

Invariant: Unchanged by a transformation of coordinates, such as the interval between two events in space-time.

Inverse beta decay: The process by which a neutrino interacts with a proton to produce a neutron and a positron.

Ion: An atom that is not electrically balanced but rather has either more electrons than protons or more protons than electrons. These therefore bear net electrical charge.

Ionization: The process whereby atoms are made into ions, by removal or addition of electrons. Such a process often occurs as a result of excitation of atoms into an energy state from which they easily lose electrons.

Ionosphere: The ionized layer of gases in Earth's atmosphere, occurring between the thermosphere (below) and the exosphere (above), between about 50 and 500 kilometers (31 to 310 miles) above the planet's surface. Within the ionosphere, ionized gases are maintained by the Sun's ultraviolet radiation. The resulting free electrons reflect long radio waves, making long-distance radio communication possible. Other planets are known to have ionospheres, including Jupiter, Mars, and Venus.

Iridium: A highly dense metallic element that is more abundant in materials of extraterrestrial origin and the Earth's core than at the Earth's surface.

Isomagnetic charts: Maps on which are traced curves, all the points of which have the same value of some magnetic property.

Isostasy: The concept that Earth's crust is in, or is trying to achieve, flotational equilibrium by buoyantly floating on denser mantle rocks beneath.

Isotope: Atoms with the same number of protons in the nucleus but with differing numbers of neutrons; a particular element will generally have several different isotopes occurring naturally.

Isotropic: Having properties that are the same in all directions; the opposite is anisotropic—having properties that vary with direction.

Jupiter-like planets: Giant gaseous planets with about the same composition as that of Jupiter, ranging from about one-tenth to ten times its mass (thirty to three thousand Earth masses).

K-band: A radio frequency range of about 11 to 15 gigahertz. *See also* Hertz.

Kamacite: A form of ferritic iron containing up to 7.5 percent nickel in solid solution.

Kelvin: A unit of temperature on the kelvin temperature scale, which begins at absolute zero (–273.15° Celsius). One unit kelvin is equal to 1° Celsius. The kelvin scale is particularly suited to scientific (especially astronomical) measurement. *See also* Absolute temperature scale.

Kepler's laws of motion: Three laws of motion discovered by Johannes Kepler and published by him in 1618-1619: (1) Each planet moves in an ellipse around the Sun, with the Sun at one of the two foci of that ellipse. (2) A line from the Sun to the planet sweeps out equal areas in equal times. (3) The square of the period of a planet's orbit is proportional to the cube of its mean distance from the Sun. *See also* Angular momentum.

Kilogram: A metric unit of mass, the equivalent of 1,000 grams. On Earth, a kilogram of mass weighs about 2.205 pounds.

Kilometer: A metric unit of distance, the equivalent of 1,000 meters or approximately 0.62 mile.

Kinetic energy: The energy of motion.

Kinetic pressure: The average force per unit area produced by movement of atoms and molecules.

KREEP-norite: A rock formation occupying small regions of the lunar surface, caused by the release of lava from liquid pockets of deep lunar crust as a result of meteoritic impacts; called KREEP because it is enriched in potassium (K), rare-Earth elements (REE), and phosphorus (P).

Kuiper Belt: A disk of icy, rocky objects that lies beyond Neptune, approximately 30 to 1,000 astronomical units from the Sun.

Lagrangian points: Stable points in the orbit of an intermediate body around a larger body where small particles may accumulate; these are sometimes referred to as L4 and L5.

Lapse rate: The rate at which temperature changes with altitude.

Laser: Originally an acronym for "light amplification by stimulated emission of radiation." A beam of infrared, visible, ultraviolet, or shorter-wavelength radiation produced by using electromagnetic radiation to excite the electrons in a suitable material to a higher energy level around their atomic nuclei. These electrons are then stimulated in such a fashion that they jump back down to their normal energy levels. When they do, they emit a stream of "coherent" radiation: photons with the same wavelength, direction, and phase as the originating radiation. This results in a narrow, intense beam of light (or nonvisible radiation), which bounces off a reflector and directly back to the propagating material, where the process is repeated and thus the laser is maintained. Laser technology has a vast range of applications in telecommunications, medicine, and astronomical measurements.

Laser ranging: A technique whereby scientists at two different Earth stations can determine, very precisely, their distance from each other by bouncing a laser beam off a satellite retroreflector. The time it takes to receive an "echo" allows each scientist to calculate his or her distance from the satellite; knowing both distances allows calculation of the distance between the two points on Earth. Over time, these measurements are repeated; changes in the distance between the two Earth locations are noted, providing much information about crustal movements and the likelihood of earthquakes.

Laser reflector: An optical device off which a laser beam can "bounce" or reflect; used to measure (usually great) distances by determining time of flight of the laser beam from original emission to final reception after reflection from the laser reflector on a body whose distance is to be determined in this way, recognizing that the laser beam travels at the speed of light (3×10^8 meters per second).

Latitude: The angular distance from a specified plane of reference. On Earth, the angular distance north or south of the equatorial plane; in the solar system, the angular distance of a celestial body above or below the ecliptic plane. *See also* Longitude.

Launch site: A location housing a facility designed to handle preparations for launch as well as the launch itself.

Lava: Molten magma from the interior of a planet extruded through cracks or holes in the planet's surface.

Leptons: A class of elementary particles, including electrons, neutrinos, and their antiparticles, which are not affected by the strong nuclear force that binds protons and neutrons.

Light, speed of. *See* Velocity of light.

Light curve: A graph that indicates how the brightness of a star changes with time.

Light-year: A unit of distance equal to the distance that light travels in a vacuum in one year. At a speed of approximately 300,000 kilometers (186,000 miles) per second, a light-year is about 10 trillion kilometers.

Limb: The outer edge of the visible disk of the Sun, Moon, a planet, or another celestial body.

Liquid metallic hydrogen: Hydrogen that behaves as a metal under pressure of about 5 million Earth atmospheres.

Lithification: The conversion of loose mineral materials into rock.

Lithium: Element 3 in the Periodic Table of Elements following hydrogen and helium, produced within protostars but destroyed by nuclear reactions in stars.

Lithosphere: A rigid layer consisting of the Earth's crust and the top part of the underlying mantle, about 80 kilometers (49.6 miles) thick. The term can be used in reference to the rigid outer part of any planet.

Longitude: The angular distance from a specified plane perpendicular to the latitude reference plane. On Earth, the angular distance east or west of the prime meridian, which runs from pole to pole through the Royal Observatory at Greenwhich, England.

Look angle: Angular limits of vision.

Low-Earth orbit: Generally, any orbit at an altitude of about 300 kilometers (186 miles) or less. Such an orbit has a period (time required to complete one orbit around Earth) of 90 minutes or less.

Lowell bands: Dark areas on the periphery of the Martian polar caps in summer.

Luminosity: The rate at which a star radiates electromagnetic energy, usually expressed in joules per second or in watts; the *intrinsic brightness* or light output of a star, as distinct from its *apparent brightness*.

Lunar day: The time it takes the Moon to complete one rotation on its axis. Relative to distant stars, it takes approximately 27.33 Earth-days. Relative to the Sun, it takes approximately 29.53 Earth-days; this is the time from one sunrise to the next, as seen from the Moon.

Lunar (synodic) month: The period from one new moon to the next, about 29.53 days.

Mach: The ratio of the speed of a moving object to the speed of sound in the surrounding medium. At Mach 1, the speed of an aircraft equals the speed of sound.

Mafic *and* ultramafic: Rock-forming magmas that are high in dense, refractory elements such as iron and magnesium; oceanic basalts are examples of mafic rocks.

Magellanic Clouds: Two small, irregular galaxies, the two nearest galaxies outside the Milky Way, visible from the Southern Hemisphere as the Large Cloud and the Small Cloud, respectively 160,000 and 185,000 light-years away. The Magellanic Clouds have been instrumental in establishing an extragalactic distance scale.

Magma: Molten rock material formed in a planet's interior; when extruded at the surface, it becomes known as lava.

Magnetic anomalies: Distortions in the magnetic field, produced by an object such as an iron ore body.

Magnetic field: A field created by magnetic properties of an object. When an object of mass m and charge q having a speed v enters an external magnetic field B created by an external influence, it experiences a magnetic sideways deflecting force given in magnitude by the product of q, v, B, and the sine of the angle between the direction of the velocity vector and the magnetic field. The direction of that sideways deflecting force is perpendicular to both the velocity and the magnetic field vectors; positive and negative charges are deflected in opposite directions from one another when entering a magnetic field. *See also* Gravitation.

Magnetic pole: The location on Earth's surface where Earth's magnetic field is perpendicular to the surface.

Magnetic storm: Rapid changes in Earth's magnetic field as a result of the bombardment of Earth by electrically charged particles from the Sun.

Magnetic surveys: Measurements of the magnetic field at many points, on or above Earth's surface, carried out by field teams, airborne magnetometers, ships at sea, or satellites.

Magnetometer: A scientific instrument used to measure disturbances in Earth's magnetic field.

Magnetopause: The outer limit of a planet's magnetic field and the boundary of its magnetosphere.

Magnetosheath: A region of magnetic turbulence between the bow shock and the magnetopause.

Magnetosphere: The domain around a planet in which the behavior of charged particles is controlled by the planet's magnetic field and not by the Sun's.

Magnetotail: A "tail" of nearly parallel lines of magnetic force extending from a planet in the direction away from the Sun.

Magnitude: The brightness of a celestial body expressed numerically; the lower the number, the brighter the body. *See also* Absolute magnitude, Apparent magnitude.

Main sequence star: A star, such as the Sun, which produces energy mainly by a hydrogen-to-helium fusion reaction. Most stars spend the greater part of their lifetimes in this state.

Mantle: A layer of dense silicate rock that lies at depths of approximately 34 to 2,885 kilometers (21 to 1,789 miles) between the crust and outer core and comprises the majority of Earth's volume.

Mare (*pl.* maria): A large, flat area on the Moon or Mars, so named (after the Latin for "sea") because these areas appear dark, thus sealike, to the Earth observer.

Maritime satellite: A satellite designed for telecommunications by and for shipping industries. These satellites occupy geostationary orbits over oceans to transmit ship-to-shore communications and data.

Mascon: One of several concentrations of mass located beneath lunar maria, which causes a distortion in the orbit of a spacecraft around the Moon.

Maser: An acronym for "microwave amplification by stimulated emission of radiation." A device similar to a laser in which energy is generated as in a laser, but at microwave levels. A maser can exist in nature as a celestial object. Artificial masers are used to amplify weak radio signals. *See also* Laser.

Mass: The amount of matter contained within a body, which determines the amount of gravitational force it exerts. Mass is measured in units such as kilograms; it differs from weight, which is the force exerted on a mass by gravity and which therefore would be measured in units such as newtons or pounds.

Mass spectrometer: An instrument that identifies the chemical composition of a substance by separating ions by mass and charge.

Matrix: In meteorities, the fine-grained material that surrounds both chondrules and inclusions; it consists of hydrous silicate minerals, troilite, magnetite, and other lower-temperature phases.

Matter: A substance that has mass and occupies space, which along with energy is responsible for all observable phenomena.

Maunder minimum: Named for E. W. Maunder, who in 1890 discovered a period in the three-hundred-year history of sunspot observations when few sunspots were recorded. Confirmed independently in 1976 by evidence from tree rings, the Maunder minimum covers the years 1645 to 1715, a period also known as the Northern Hemisphere's "Little Ice Age."

Megahertz. *See* Hertz.

Megaparsec: A unit of measurement equaling 3.26 million light-years.

Mesosphere: The layer of Earth's atmosphere occurring above the stratosphere and below the thermosphere, from about 40 kilometers to 85 kilometers (24.8 to 52.7 miles) above sea level. This is the coldest layer of the atmosphere.

Messier number: The number of an object listed in the catalog of 103 nebulae and star clusters prepared by Charles Messier in 1784; an object is referred to as M followed by the catalog number, such as the globular star cluster M13 in Hercules.

Metabolic process: A chemical process in a living organism which provides it with the energy to function.

Metamorphism: A process by which heat and pressure applied to a rock cause it to change without causing it to melt.

Meteor: A bright streak of light in the sky, sometimes called a shooting star, produced by a meteoroid entering Earth's atmosphere at high speed and heating the air column along its path to incandescence.

Meteor shower: A large number of meteors resulting from the passage of Earth through a meteoroid stream or swarm believed to be the debris left in the orbit of a comet.

Meteorite: A metallic or stony meteoroid (or combination) that survives its passage through the atmosphere as a meteor and falls to the surface of the Earth.

Meteoritics: The study of the naturally occurring masses of matter that have fallen to the Earth's surface from outer space.

Meteoroid: A small solar system body, probably a fragment from a comet or asteroid, which produces a meteor when it enters Earth's atmosphere.

Meter: The metric unit of length, equivalent to approximately 39.37 inches, or a little more than 1 yard.

Metric system: The decimal system of weights and measures, which forms part of the Système International d'Unités. *See also* SI units.

Metric ton (or tonne): A unit of mass equal to 1,000 kilograms. On Earth, it is equivalent to 2,205 pounds, which is close to an ordinary ton, or short ton, of weight. *See also* Newton, Pound.

Microgravity: Nearly zero gravity. Microgravity exists in a space vehicle because of the minute gravitational forces exerted by objects on one another. The microgravity environment is of great importance as an ideal environment for certain types of materials processing.

Micrometeorite: A micrometeoroid that has reached Earth's surface.

Micrometeoroid: A meteoroid with a diameter of less than 0.1 millimeter. Because of their size, micrometeoroids rarely burn up but reach Earth's surface instead, as spherules or as cosmic dust particles.

Micron: A unit of measure convenient for measuring wavelengths of infrared radiation; a micron is equal to one millionth of a meter.

Micropaleontology: The study of microscopic fossils, of potential importance in exobiology as well as life sciences on Earth.

Microwaves: A form of electromagnetic radiation with wavelengths ranging between 1 millimeter and 30 centimeters, located between infrared and long-wave radio on the electromagnetic spectrum.

Milky Way: The galaxy in which our solar system is located. The Milky Way, a spiral or barred spiral galaxy, contains about 10^{11} stars and is about 100,000 light-years across. *See also* Galaxy.

Miller-Urey synthesis: The production of amino acids by repeatedly passing an electrical spark through a mixture of methane, ammonia, water vapor, and hydrogen.

Millibar: A pressure of 100 newtons per square meter.

Missing mass of the universe. *See* Dark matter.

Model: A simulation of a phenomenon that is difficult to observe or specify by direct means; models abstract from phenomena under study those qualities that the investigator perceives to be essential for understanding.

Molecular clouds: Massive, very large clouds of various molecules and dust; the breeding ground of stars in deep space.

Molecule: The smallest unit of a substance, formed by a characteristic complex of atoms joined together. The smallest unit of the substance water, for example, is a molecule formed by two hydrogen atoms and one oxygen atom.

Moon: Any natural satellite orbiting a planet, especially Earth's Moon.

Nanometer: One-thousand-millionth (one-billionth) of a meter; a unit used to express electromagnetic wavelengths.

Near-Earth space: Roughly defined as the space environment from the outer reaches of Earth's atmosphere to the path of the Moon's orbit, the area beyond which is known as deep space.

Near-infrared: A portion of the electromagnetic spectrum that lies beyond the red end of the visible spectrum and has a photon energy that is approximately the same as that typical of many chemical bonds.

Nebula: A celestial body composed of aggregated gas and dust, which may be either luminous, reflecting or emitting light under the influence of nearby stars (a reflection nebular or an emission nebula), or dark, obscuring the light of distant stars and appearing as a silhouette (an absorption nebula).

Nebular hypothesis: The concept that the solar system and all of its parts are the result of the contraction of a cloud of gas and dust.

Neumann bands: A textural pattern that is common to iron meteorites with less than 6 percent nickel content; they reflect deformational twinning paralled to trapezohedral planes in kamacite.

Neutral gas analyzer: An instrument that determines the chemical composition of an atmosphere.

Neutrino: An elementary particle of enormous penetrating power as a result of its lack of electric charge and its nearly total lack of mass. Traveling directly out from the cores of stars as a by-product of nuclear reactions, neutrinos have enormous potential as a source of information on the stars and other astrophysical phenomena.

Neutron: An elementary particle composed of quarks in such a way that it has no net electric charge. Found in atomic nuclei and is attracted to other nucleons (other neutrons and protons) by the strong nuclear force. Its mass is approximately equal to that of a proton.

Neutron stars: The smallest stars known, with diameters of about 20 kilometers (12.4 miles) and masses about two times that of the Sun, consisting of a thin iron shell enclosing a "liquid" sea of degenerate neutrons.

Although astrophysicists lack a full understanding of these objects, neutron stars are thought to originate from stars much more massive than the Sun which explode as supernovae. Much mass is lost in the process, and what remains as the neutron star may spin rapidly and be observable as pulsars.

New astronomy: A term used collectively to refer to the areas of astronomy (such as gamma-ray astronomy, infrared astronomy, and X-ray astronomy) investigating electromagnetic emissions by celestial phenomena observable from satellites in orbit around the Earth or Sun.

New General Catalog: A catalog—in full, the *New General Catalog of Nebulae and Clusters of Stars*—compiled in 1888 by Johan Ludwig Emil Dreyer, which lists many nebulae, star clusters, and galaxies. Frequently the objects listed in the catalog are identified by their NGC numbers; for example, the globular star cluster Omega Centauri is NGC 5139.

Newton: The basic metric or SI (Système Internationale) unit of measure for force or weight; also used to measure thrust. Named after Sir Isaac Newton.

North point: That intersection of the celestial meridian and astronomical horizon lying due north.

Northern lights. *See* Aurora.

Nova: A star that emits a sudden burst of visible light and other forms of electromagnetic radiation, and quickly (over months or years) returns to its former brightness. *See also* Supernova.

Nuclear burning: The process by which the nuclei of light elements, under conditions of extreme temperature and density, are converted into the nuclei of heavier elements through fusion.

Nuclear energy: Energy that is released as a result of interactions between elementary particles and atomic nuclei.

Nuclear reaction: A very high-energy process in which nuclei break apart (or fission), if very massive, or join together (or fusion) if very light, causing the emission of heat and light.

Nucleosynthesis: The process by which heavier elements are produced from hydrogen and helium in stars.

Nucleus: The central part of something. In galaxies, the central region or core; in spiral galaxies, sometimes called the central bulge; in comets, the solid body of the comet; in atoms, the central massive part consisting of protons and neutrons about which the electrons orbit.

Oblate spheroid: A nearly spherical shape that is flattened at the poles and bulges at the equator.

Occultation: An eclipse of any astronomical object other than the Sun or the Moon caused by the Moon or any planet, satellite, or asteroid.

Octahedrites: Iron meteorites that usually contain between 6 and 16 percent nickel.

Olivine: A silicate mineral of magnesium and iron that is common in some igneous chondritic meteorites.

Oort Cloud: A vast region of billions to trillions of comets that surrounds the solar system and is located tens of thousands of astronomical units from the Sun.

Open universe: A universe that expands forever and will eventually lose all usable energy.

Opposition: The alignment of Sun, Earth, and a superior planet (one whose orbit is farther from the Sun than Earth's) in a straight line; that is, the superior planet appears in the sky at 180° celestial longitude from the Sun. In this position, the planet is closest to Earth and therefore most easily observed by ground-based instruments.

Optical radiation: Visible light; a range of wavelengths that form a small section of a larger range of wavelengths called the electromagnetic spectrum.

Optical window: The region of the electromagnetic spectrum (295 to 1,100 nanometers) that is passed by the atmosphere and that is easily manipulated by lenses and mirrors; "visible" light (400 to 700 nanometers) lies near the center of this window.

Orbit: The path traced out by one body as it moves around another. The distinguishing characteristics of an orbit are called its orbital parameters and include apoapsis, periapsis, inclination, eccentricity, and period. All closed orbits trace out an ellipse, of which the body orbited lies at one of the two foci. *See also* Apoapsis, Circular orbit, Eccentricity, Ellipse, Elliptical orbit, Equatorial orbit, Geocentric orbit, Geostationary orbit, Geosynchronous orbit, Heliocentric orbit, Orbital inclination, Parabolic orbit, Parking orbit, Periapsis, Period, Polar orbit, Prograde orbit, Retrograde orbit, Synchronous rotation (or orbit), Transfer orbit.

Orbital eccentricity: A measure of the elongation of an elliptical orbit, ranging from zero (for a circular orbit) to one (for an orbit that approaches an open-ended parabola).

Orbital inclination: The angle formed between the orbital plane of a satellite and the equatorial plane of the object orbited.

Orbital period: The time required for one celestial object to execute a complete revolution around another object.

Ordinary chondrites: Chondrule-bearing stony meteorites not containing carbonaceous compounds.

Organic molecules: Molecules including the elements carbon, hydrogen, and oxygen.

Oscillating universe: A theoretical type of closed universe in which a big bang and subsequent recompression of all matter periodically occur forever.

Outer core: A zone of the Earth's interior, located at depths of approximately 2,885 to 5,144 kilometers (1,789 to 3,189 miles), that is in a liquid state and consists of iron sulfides and iron oxides.

Outer space: All space beyond Earth's atmosphere. *See also* Deep space.

Outgassing: The process by which trapped gases in a planet leak out gradually over time to form the planetary atmosphere and oceans.

Ozone layer: The thin layer of Earth's atmosphere, located between 12 and 50 kilometers (7.44 to 31 miles) above Earth's surface (in the stratosphere), in which ozone (O_3) is found in its greatest concentrations. This layer, which absorbs most of the ultraviolet radiation entering the atmosphere, forms a protective blanket around the planet, shielding it from excess radiation.

P waves. *See* Primary (P) waves.

Paleomagnetism: The study of the record of remanent or fossil magnetism in rocks, indicative of past states of Earth's magnetic field and very useful in determining secular variation, sea-floor spreading, and past locations of continents.

Panspermia: A theory proposed by chemist Svante Arrhenius in 1906, and later modified by Sir Fred Hoyle, which holds that organic molecules (hence the beginnings of life) were transported to Earth by means of comets.

Parabolic orbit: An orbit that describes a parabola around the object orbited and hence escapes from the gravitational field of that object.

Parallax: The apparent angular displacement of a star as seen from opposite sides of the Earth's orbit around the Sun. Measuring this angle (and knowing the size of the Earth's orbit) allows the distance to the star to be calculated.

Parking orbit: An interim orbit around a celestial body between launch and injection into another orbit or into a trajectory toward another destination.

Parsec: A unit for measuring astronomical distances equivalent to 3.26 light-years; the distance at which the stellar parallax is 1 second of arc.

Partial lunar eclipse: An eclipse in which the Moon passes completely or partially into the penumbra of the Earth but does not enter (or entirely enter) the umbra.

Partial melting: Melting of some minerals in a rock but not others, resulting in a magma concentrated in some elements and depleted in others as compared to the original unmelted rock.

Partial solar eclipse: An eclipse in which part but not all of the Sun is covered by the Moon.

Particle. *See* Elementary particles.

Path of the Sun: The apparent motion of the Sun as it tracks across the sky.

Pauli exclusion principle: The principle that no two particles of the same type can occupy precisely the same quantum state; it is obeyed by baryons, leptons, and quarks, but not by photons or mesons.

Pedestal crater: A crater that has assumed the shape of a pedestal as a result of the wind's unique shaping processes.

Penumbra: A region of partial shadow where some but not all parts of the source of illumination are obscured.

Perfect spheroid: A three-dimensional body that has the same circumference regardless of the direction by which it is measured; that is, it is perfectly "round."

Periapsis: The point in one object's orbit around another at which the orbiting object is closest to the object being orbited.

Pericynthion: The point in an object's orbit around the Moon at which it is closest to the Moon.

Peridotite: A silicate igneous rock consisting largely of the mineral olivine.

Perigee: The point in an object's orbit around Earth at which it is closest to Earth.

Perihelion: The point in a solar orbit at which the orbiting object is closest to the Sun.

Perilune: Pericynthion of an artificial satellite.

Period: The time span between repetitions of a cyclic event. An *orbital period* is the time required for a satellite or moon to make one complete orbit around a planet, a moon, the Sun, or another celestial body.

Permafrost: Permanently frozen soil which is laced with water ice.

Permeability: The property or capacity of porous geological materials to transmit fluids; it indicates the relative ease of fluid flow through a medium.

Perturbation: The act of altering the orbital course (direction or speed) of satellite or planetary orbits usually initiated by gravitational effects from or collision with another object.

Petrography: The description and systematic classification of rocks.

Photochemistry: Chemical reactions caused by the action of strong ultraviolet light, which excites or dissociates some compounds and leads to the formation of others.

Photometer: An instrument that measures the brightness of a light source.

Photometry: The technique of measuring the brightness of light sources.

Photomultiplier: An instrument for increasing the apparent brightness or strength of a source of light by means of secondary excitation of electrons; effectively, a light (or other radiation) amplifier.

Photon: A particle of electromagnetic radiation possesing energy and momentum but neither rest mass nor charge.

Photopolarimeter: An instrument used to measure the brightness and polarization of light.

Photosphere: The region of the Sun that separates its exterior (the chromosphere and corona) from its interior, forming the boundary between the transparent and opaque gases. The photosphere appears as the bright central disk from Earth, and it is the source of most of the Sun's light.

Photosynthesis: The utilization of carbon dioxide and water by chlorophyll-containing organisms in the presence of sunlight to metabolically produce carbohydrates used by the plant for food; oxygen is a by-product in the photosynthesis process.

Photovoltaic cell: A solid-state energy device that converts sunlight into electricity.

Pitch, roll, and yaw: Movements that a spacecraft undergoes as a result of launch or other stresses. Pitch is up-down movement; roll is longitudinal rotation; yaw is side-to-side movement.

Pixel: A small unit arranged with others in a two-dimensional array which contains a discrete portion of an image (as on a television screen) or an electrical charge (as on a charge coupled device). Together, these pixels form an image or other meaningful information.

Planck length, *l*: The Planck length l_p is defined by $[Gh/(2pc^3)]^{0.5}$ and is approximately equal to 1.6×10^{-35} meters. Combining gravitation's G with quantum mechanics' h, it marks the distance at which quantum effects dominate.

Planck's constant, *h*: Named after the German physicist Max Planck, h (6.63×10^{-34} joule-seconds) first arose in modern physics in an attempt to explain blackbody radiation characteristics. It is a quantum of "action," having units of angular momentum. Among its many fundamental aspects in quantum physics, it is the constant of proportionality between the energy and frequency of a photon of light.

Plane of the ecliptic: The plane of the Earth's orbit around the Sun.

Planet: A nonluminous natural celestial body that orbits the Sun (or another star) and is not categorized as an asteroid or comet. There are eight known planets in the solar system: Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, and Neptune in increasing distance from the Sun. Pluto lost it planetary status as a result of a redefinition of "planet" adopted by the International Astronomical Union (IAU) in 2006.

Planetesimal: A small solid body (from less than a millimeter up to hundreds of kilometers) that accreted in the solar nebula during the formation of the solar system. Many of these bodies grew into protoplanets and planets, while others remain as meteoroids and asteroids.

Plasma: Ionized gas, consisting of roughly equal numbers of free electrons and positive ions. Plasma occurs in stars, nebulae, and interplanetary, interstellar, and intergalactic space. It is considered the fourth state of matter, along with solid, liquid, and gas.

Plasma sheath: The definite outer boundary of Earth's ionosphere.

Plate tectonics: The theory that the crust and upper mantle of the Earth are divided into a number of moving plates about 100 kilometers (62 miles) thick that converge at trench sites and diverge at oceanic ridges.

Polar orbit: An orbit in which a satellite passes over a planet's or moon's poles.

Polarimeter: An instrument for measuring the degree to which electromagnetic radiation is polarized.

Pole: One of two points on the surface of a planet where it is intersected by its axis of rotation. In a magnetic field, one of two or more points of concentration of the lines of magnetic force.

Polygonal ground: The distinctive geological formation caused by the repetitive freezing and thawing of permafrost.

Polymorphism: The characteristic of a mineral to crystallize into more than one form.

Porphyritic chondrules: Generally small spherules (1 to 5 millimeters), often with crystals of the minerals olivine and pyroxene set into a glass matrix.

Posigrade: Moving in the direction of travel.

Positron: The antiparticle of an electron.

Potential energy: Energy due to an object's location in some sort of field, such as gravitational potential energy due to a gravitational field or electrical potential energy due to an electric field.

Pound: An Imperial unit used to measure force, weight, and thrust in the United States and some other English-speaking nations.

Precession: A type of motion that occurs in a rotating body in response to torque: A planet or other rotating body orbiting around the Sun will slowly turn its rotation axis in such a way that, over a long period of time, the rotating each of the body's two poles describes a circle in space. The fact that the Earth precesses means that adjustments must be made in the direction in which astronomers look to observe stars and other celestial phenomena. Earth's period of precession is approximately 25,800 years.

Primary minerals: Those minerals formed when magma crystallizes.

Primary (P) waves: The fastest elastic wave generated by an earthquake or artificial energy source; basically an acoustic or shock wave that compresses and stretches material in its path.

Primordial solar nebula: An interstellar cloud of gas and dust that condensed under gravity to form the Sun, the planets and their satellites, asteroids, comets, and meteoroids some 4.6 billion years ago.

Principle of equivalence: The rule that, in a limited region of space-time, the effects of the acceleration of a given frame of reference are not distinguishable from those of a gravitational field.

Probe. *See* Deep space probe.

Prograde orbit: An orbit that moves in the same direction as the rotation of the body orbited.

Prominences: Arches of glowing gases above the Sun's photosphere often seen in the vicinity of sunspots.

Proper motion: The motion of a star across the line of sight.

Protein: A high-molecular-weight compound that is a long chain or aggregate of amino acids joined by hydrogen bonds.

Proto solar system (solar nebula): The cloud of gas and dust that separated from a larger cloud and eventually collapsed to form the Sun in the center and the planets in the outer portions of the cloud.

Proton: The constituent particle in the nucleus made up of quarks in such a way that it has a net positive electric charge. It interacts with other protons and neutrons inside a nucleus via the strong nuclear force, interacts with electrons and protons (at distances greater than nuclear dimensions) via the Coulomb electric force, and decays according to the weak nuclear force. Chemically speaking, the number of protons in the nucleus of an atom determines which chemical element it is.

Protoplanet: A stage of planet formation in which a large precursor of a planet gravitationally contracts, perhaps attracting more material or maybe losing part of a gaseous envelope.

Protostar: The stage just before actual star formation, when a vast cloud of matter has coalesced to about 1 percent of the cloud's dispersed volume, causing nuclear reactions to begin.

Protosun: A stage of star formation in which gravitational contraction shrinks and heats the developing star, before the interior becomes hot enough to initiate nuclear fusion reactions.

Pulsar: A rapidly spinning neutron star that emits a narrow beam of electromagnetic radiation in the form of visible light, radio waves, X rays, and gamma rays.

Pyroxene: A silicate mineral of magnesium, iron, and sometimes calcium which contains more silicon than is present in olivine.

Quantum theory: An important theory in physics that associates waves with particles; particle momentum is h divided by the particle's wavelength, while energy is h times the particle's frequency; light has frequency and wavelength (thus, energy and momentum), and so acts like a particle and a wave.

Quarks: Subatomic particles hypothesized to form electrons, protons, neutrons, and their antiparticles, characterized by electric charge, "flavor," and "color." The forces required to break elementary particles into their component quarks is so great that quarks do not exist as free particles in nature, although it is thought that neutron stars may consist of a "quark soup" within a solid iron shell.

Quasar: An acronym for "quasi-stellar radio source" or "quasi-stellar object." An object continuously releasing a tremendous amount of energy, equivalent to the output of between one million and 100 trillion suns, which includes virtually all kinds of electromagnetic radiation (gamma, X, ultraviolet, optical, infrared, microwave, and radio) from a very small volume of space about the size of the solar system. As far as is known, all objects satisfying these criteria are located in the nuclei of galaxies; it is thought that they may be associated with supermassive black holes at the centers of galaxies. Discovered in 1963, the first quasar caused much excitement among astronomers, and these phenomena continue to be among the most fascinating and mysterious in the universe.

Radar: An acronym for "radio detection and ranging." A means of locating and determining the distance of objects by bouncing radio waves off them and measuring the time required to receive the echo.

Radial velocity: Movement in the line of sight, toward or away from the observer.

Radiant: The point in the sky from which a meteor shower seems to emanate, whose associated constellation provides the name for a given shower.

Radiating pyroxene chondrules: Generally small spherules (1 to 5 millimeters) composed of excentroradial pyroxene crystals, often resembling a fanlike growth pattern.

Radiation. *See* Electromagnetic radiation.

Radiation (of stars): Stars radiate energy in all portions of the electromagnetic spectrum, whereas one band of wavelengths predominates according to the star's photospheric temperature.

Radiative equilibrium: When radiation becomes the predominant, balanced mechanism for carrying heat away from the interior of a star.

Radio astronomy: The branch of astronomy that examines the radio emissions of celestial objects. Because radio radiation, along with visible radiation, can penetrate Earth's atmosphere, radio receivers have provided much of the data detectable by ground-based, as well as space-based, instruments. Radio emissions also form a significant portion of certain celestial phenomena, such as radio galaxies, quasars, and pulsars.

Radio noise: Any body that generates radio waves or oscillations of a random nature over the radio frequencies, from millimeters to several thousand meters; often heard as a hissing sound in radio receivers.

Radio spectrum: Those frequencies generally ranging from 1 centimeter to about 30,000 centimeters.

Radio telescope: A telescope designed to gather radio waves from extraterrestrial sources.

Radioactive decay: The conversion of one element into another by the emission of charged particles from an atom's nucleus.

Radioactivity: The process by which an unstable atomic nucleus spontaneously emits a particle (or particles) and changes into the nucleus of another atom.

Radiogenic heating: Heating caused by the decay of radioactive materials in a planetary body; energy released during the decay produces the heat.

Rampart crater: A type of crater found most often on Mars and produced by some subsurface shaping mechanism that causes a unique, rampart-type wall formation.

Real time: Referring to the transmission of signals or other data at the same time that they are used.

Reconnaissance satellite: A satellite that gathers information about enemy military installations.

Red dwarf star: A relatively small, cool star having low luminosity.

Red giants: Stars with surface temperatures less than 4,700 kelvins and between 10 and 1,000 times the diameter of the Sun, which emit mostly red and infrared light, and have huge surface areas.

Red Planet: Mars, so named because of its color as seen through Earth-based telescopes.

Redshift: The apparent lengthening of electromagnetic wavelengths issuing from an object as a result of the object's movement away from the observer. As a result, the spectral lines in the spectra of such an object will shift toward (or beyond) the red end of the visible spectrum. *See also* Doppler effect.

Reentry: The return of a spacecraft into Earth's atmosphere.

Reflecting telescope: An optical telescope that uses a mirror or mirrors to gather and focus light from the object observed. These telescopes are widely used for Earth-based as well as space-based optical astronomy. *See also* Cassegrain telescope, Refracting telescope.

Reflection: The "bounce" of wave energy off a boundary that marks a change in the material; it must be noted that at an interface between two different media, wave energy can be reflected to a certain degree back into the original medium, and what is not reflected off the interface is refracted across the interface into the other medium.

Reflection nebula: A cloud of dust, usually bluish in color, visible by virtue of light from nearby stars scattered by the dust grains.

Reflectivity: The amount of light reflected from a body.

Refracting telescope: An optical telescope that uses a lens to magnify and focus light from the object observed. Refracting telescopes were used by the earliest astronomers. When reflecting telescopes were perfected in the twentieth century, refracting telescopes became less important in astronomy, although they are still widely used for guided and amateur observations. *See also* Reflecting telescope.

Refraction: The change in direction of a wave path upon crossing a boundary, resulting from a change in the material light enters into and the different velocity of the wave on either side of the boundary.

Refractory: Refers to substances that melt or boil at relatively high temperatures and, conversely, condense from liquids or gas at low temperatures.

Refractory (siderophile) elements: Elements least likely to be driven off by heating; the last elements to be melted as a rock is heated to form magma.

Regolith: The layer of soil and rock fragments just above the solid planetary crust.

Relativity: The physical law, first proposed by Albert Einstein, that states that measurements of time and space are dependent upon the frame of reference in which they are measured. The general theory of relativity applies this law to gravity and mass; the special theory of relativity applies it to the propagation of electric and magnetic phenomena in space and time.

Remote sensing: Acquiring data at a distance by electronic or mechanical means.

Resolution: The degree to which a photographic or other imaging system, or the image produced, clearly distinguishes objects of a certain size. In a photograph with a resolution of 200 meters, for example, the smallest distinguishable objects are 200 meters across.

Retrograde orbit: An orbit that moves opposite to the rotational direction of the body orbited.

Revolution: One complete orbit of one body around another body, such as a planet around the Sun or a natural or artificial satellite around a planet.

Right ascension: A coordinate for measuring the east-west positions of celestial bodies; the angle measured eastward along the celestial equator from the vernal equinox to the hour circle passing through a body.

Rille: A long, narrow valley on a moon or planet.

Ring resonance: The gravitational interaction between a ringed planet, its moons, and the particles in the rings; the principal effect of the resonance is formation of discrete rings and gaps.

Robotics: The development, construction, and use of computerized machines to assist or replace humans in a variety of tasks requiring precise "hand-eye" coordination.

Roche limit: Named for Édouard Roche, who discovered it in 1848, the minimum distance from a planet at which a natural satellite can form by accretion: roughly 2.44 times the planet's radius. Within this limit an existing satellite will be torn apart by gravitational stresses. Saturn's rings, which lie within the planet's Roche limit, may be the remnants of a former moon.

Roll. *See* Pitch, roll, and yaw.

Rotate: To spin around an axis.

RR Lyrae stars: A class of regular pulsating variable stars, having a period of about half a day and an average luminosity of about one hundred times that of the Sun.

Saros cycle: A period of 223 lunar months, after which the pattern of eclipses repeats.

Satellite: Any body that orbits another of larger mass, usually a planet. Satellites include moons, the small bodies that form planetary rings, and human-made robotic spacecraft. *See also* Artificial satellite.

Scarp: A vertical or near-vertical cliff, often extending for many kilometers.

Schwarzschild radius: The radius of a mass, such as a collapsing star, at which it becomes a black hole—that is, at which its gravitational field will not allow light to escape. The length of this radius depends on the body's mass, and the formula for calculating it was established by Karl Schwarzschild in 1916.

Scintillation: Light emitted when high-energy radiation is absorbed by matter then re-emitted at lower energies.

Scintillation counter: A device that detects atomic particles and generates an electrical current proportional to the energy of the particle.

Secondary crater: A crater resulting from impact of material thrown out of a primary impact crater.

Secondary (S) wave: A transverse type of earthquake wave, slower than a primary wave, which will not travel in a liquid.

Secular variation: A change in the magnetic pole position on the Earth's surface over hundreds of years.

Sedimentary rocks: Rocks that are formed by the deposition of layers of sediment (the weathered and eroded debris from preexisting rocks).

Seismic activity: Any movement in the outer layer of a planet or moon.

Seismic waves: Elastic oscillatory disturbances spreading outward from an earthquake or human-made explosion; they provide the most important data about Earth's interior.

Seismometer: A sensitive instrument that measures movements in the outer layer of a planet or moon. The graphs produced by this instrument can be interpreted to determine the magnitude and intensity of seismic activity.

Selenography: The study of lunar surface features; the counterpart of geography on Earth.

Selenology: The study of the Moon, analogous to geology on Earth.

Self-exciting dynamo: Also called a self-sustaining dynamo, a model of the earth's outer core in which the magnetic field produced by convection is in the same direction as the field through which the motion occurs.

Seyfert galaxy: A spiral galaxy with an active core that is similar to a quasi-stellar object.

Shear (S) waves: Seismic waves transmitted by an alternating series of sideways movements in a solid; they cannot be transmitted through liquids or gases.

Shepherd satellites: Also called shepherd moons, the tiny satellites responsible for gravitationally restraining ring particles in their defined rings.

Shock wave: A zone of compression and heating of matter traveling faster than the speed of sound in the matter.

SI units: The collective units of measurement used in the Système International d'Unités, the system of measurement most widely used by scientists. Its fundamental, or *base* units are seven: the meter (the base unit of length), kilogram (mass), second (time), ampere (electric current), kelvin (temperature), mole (amount of substance), and candela (luminosity). From these seven base units, other units are derived, which are multiples, fractions, or powers of the base units such as the kilometer (1 meter × 10^3) and the square meter (the unit of area). Further derived units are derived from combinations of the base units and have their own names: hertz (the unit of frequency, which is cycles per second), newton (force or thrust, kilogram-meters per second squared), pascal (pressure, newtons per square meter), joule (energy, newton-meter), watt (power, joules per second), coulomb (quantity of electricity, the ampere-second), volt (electric potential, watts per ampere), farad (capacitance, or the ability to store energy, coulombs per volt), and ohm (electrical resistance, volts per ampere). In the United States, the base SI units are coming into increasing use. Some measures, however, remain more familiarly rendered by English units of measure, even in scientific use: It is common, for example, to refer to rocket thrust in pounds rather than newtons, and atmospheric pressure is often measured in pounds per square inch (or bars and millibars).

Sidereal period: The time for a planet or satellite to make one complete rotation on its axis, or one complete revolution around its orbit relative to distant stars.

Siderophile elements. *See* Refractory elements.

Silicate: A class of mineral whose lattice structure includes one silicon atom surrounded by four oxygen atoms at the vertices of a tetrahedron.

Singularity: In space-time, a location at which matter and energy are compressed down to a single point in space with infinite density. The initial condition of the big-bang universe, when space-time curvature and mass-energy density were infinite.

Sinuous rille: A riverlike channel produced by lava flowing across the lunar surface.

Smokers: Undersea vents in the active rift areas, emitting large amounts of superheated water and dissolved minerals from deep inside Earth.

Smooth plain: A formation that is relatively flat, with a sparsely cratered surface.

Snell's law: A statement of the fact that refraction of waves across a boundary will occur such that the ratio of the two velocities of the material on either side of the boundary is equal to the ratio of the sines of the two angles on either side of the boundary formed by the ray path and a line perpendicular to the boundary.

Solar array: An assembly of solar cells, as on a solar panel extending from a satellite.

Solar cell: A photovoltaic device that converts solar energy directly into electricity.

Solar constant: The amount of solar energy received by a square meter in one second just above Earth's atmosphere; approximately 1,370 watts per square meter per second.

Solar cycle: A period of approximately eleven years during which the number of sunspots visible near the Sun's equator increases to a maximum and then decreases. Other solar activity follows the solar cycle. *See also* Sunspots.

Solar flare: A large eruption of charged particles and electromagnetic radiation ejected from the Sun's surface (in the low corona and upper chromosphere) and lasting from a few minutes to several hours. Solar flare activity affects radio transmission on Earth and can produce auroras in Earth's atmosphere.

Solar mass: A unit equivalent to the mass of the Sun, or 1.989×10^{30} kilograms. Masses of other stars are often expressed in terms of solar masses.

Solar nebula: A cloud of mostly hydrogen and helium gas from which the Sun and planets are believed to have formed between 4.5 and 5.0 billion years ago.

Solar radiation: The radiation emitted by the Sun.

Solar system: The Sun and everything that orbits it, including planets and their satellites, plus numerous comets, asteroids, meteoroids, and other objects. *See also* Planet.

Solar ultraviolet radiation: Electromagnetic radiation emitted by the Sun in the spectral interval between approximately 90 and 400 nanometers.

Solar wind: The hot ionized gases, or plasma, that escape the Sun's gravitational field and flow in spirals outward at about 200 to 900 kilometers (124 to 558 miles) per second. It consists primarily of free protons, electrons, and alpha particles escaping from the Sun's corona. *See also* Stellar wind.

Sonar: An acronym for sound navigation ranging. A system for bouncing sonic and supersonic waves off a submerged object in order to determine its distance.

Sounding sensor: A sonarlike device that probes an atmosphere to detect data about temperature, moisture, and other conditions.

Space age: The age of space exploration, whose beginning is generally dated from October 4, 1957, the day on which the first artificial satellite, Sputnik 1, was launched into Earth orbit.

Space telescope: Any astronomical telescope that operates in space rather than on Earth.

Space-time: A four-dimensional coordinate system consisting of the three spatial dimensions and time as a fourth crewed.

Spacecraft: Any self-contained, crewed or uncrewed, space vehicle; more specifically, a deep space probe.

Specific gravity: The ratio of the density of a substance to the density of water.

Specific heat: The number of calories of heat required to raise the temperature of one gram of a substance 1° Celsius.

Spectral class: A system of classifying stars based on the pattern of absorption lines appearing in their spectra. Physically it indicates the photospheric temperature of the stars; in order of decreasing temperature the sequence runs O, B, A, F, G, K, M, L, T.

Spectrograph: A type of spectrometer that splits light into its component wavelengths and records the separated wavelengths photographically or by means of a charge coupled device.

Spectrometer: An instrument for obtaining and measuring a spectrum.

Spectroscope: An instrument that spreads electromagnetic radiation into its component wavelengths.

Spectroscopy: The science of breaking up light into its various components and studying them.

Spectrum: The entire range of electromagnetic radiation from long-wavelength radio waves to short-wavelength gamma rays; also, a limited range of wavelengths in which an instrument separates the component wavelengths or frequencies.

Speed of light. *See* Velocity of light.

Spherules: Rounded glass particles, probably formed by rapid cooling of molten material.

Spin: The rotation of a body about an axis through itself.

Spin axis: The line around which a body rotates.

Spiral arm: A dense region of heavy star formation in the disk of a spiral galaxy.

Spiral galaxy: A galaxy consisting of a bulge of gas and stars at the center, around which "arms" of stars and other celestial bodies, matter, and radiation rotate in a spiral fashion.

Sputtering: The bombardment of solid surfaces by high-energy particles, such that atoms and molecular fragments are eroded from the surface and the surface chemistry is altered.

Stagnation point: The point at which the repulsive force of a planet's magnetic field balances the pressure from the solar wind.

Star: A light-emitting body composed principally of hydrogen, whose heat and light are sustained by nuclear fusion, with a mass that is about one-tenth to one hundred times that of the Sun, a surface temperature ranging from about 3,000 to 50,000 kelvins, and an interior temperature ranging from about one million to hundreds of millions of kelvins.

Star color and temperature: Light energy from cool stars is emitted mostly at the low-energy red and infrared end of the spectrum, whereas hot stars are blue-white and produce more high-energy blue and ultraviolet light.

Star tracker: An electronic device programmed to detect and lock onto a celestial body, such as the star Canopus, to provide a spacecraft with a fixed point of reference for purposes of navigation.

Steady state theory: A model of the universe proposed by Hermann Bondi, Thomas Gold, and Fred Hoyle, which posits that the density of matter in the universe remains constant in an expanding universe, with new matter spontaneously created in the space between galaxies as they move apart. The theory is now generally considered obsolete because it does not explain the presence of the cosmic background radiation. *See also* Big bang theory, Cosmic background radiation, Dark energy, Dark matter.

Stefan-Boltzmann law: A relationship between the temperature of a "blackbody" and the rate at which it radiates energy.

Stellar core: The high-temperature, high-density region at the centers of stars where the nuclear reactions that form the heavy elements occur; no more than a few percent of the entire stellar radius.

Stellar evolution: The theory that describes the changes that occur in the internal structure of stars during their lifetimes.

Stellar luminosity: Energy emitted by per second by a star, commonly measured in joules per second (watts).

Stellar photosphere: The deepest layer of a star subject to direct observation, where the absorption lines caused by individual elements form; it lies at the top of the star's interior and the bottom of its atmosphere.

Stellar populations: A differentiation between stars of different age determined from their metal abundances and distribution.

Stellar wind: The ionized gases, or plasma, that flow out from stars at high speeds, composed mainly of free protons and electrons. *See also* Solar wind.

Stratosphere: The layer of Earth's atmosphere between the troposphere and the mesosphere, extending from about 15 to 50 kilometers (9.3 to 31 miles) above Earth's surface, roughly coinciding with the ozone layer, which absorbs the Sun's ultraviolet radiation and heats the stratosphere from a low of about –60° Celsius (213 kelvins) at the bottom to about 0° Celsius (273 kelvins) at its top. There is no meteorological activity or vertical air movement in this region of the atmosphere.

Strong nuclear force: The strongest of the four known fundamental forces in the universe; it acts over a very short range, essentially only within the nucleus itself. Details of the strong nuclear force are described by the standard model.

Sublimation: The transformation of a solid directly into a gas or a gas directly into a solid, passing through no liquid stage.

Subsatellite: A satellite carried into orbit by another satellite. Also, a satellite of a moon.

Sunspots: Dark spots that appear on the Sun's surface in cycles, their numbers increasing and decreasing with the approximately eleven-year solar cycle. These dark spots are about 500 kelvins cooler than a surrounding lighter area of the spot, which in turn is another 500 kelvins cooler than the surrounding photosphere. These cooler regions result from strong magnetic fields.

Supercluster: A gravitationally bound accumulation of galaxy clusters; most galaxies belong to clusters, with clusters of galaxies belonging to superclusters.

Superfluid: A fluid that flows without viscosity, as exemplified by liquid helium below its Lambda temperature.

Supergiants: The very largest stars. Red supergiants are up to one thousand times the size of the Sun.

Superimpositions: Craters that are formed within other craters, such as those formed when a meteoroid hits inside or on the walls of an existing crater.

Supernova: A violent explosion that tears a star apart. It can occur in a massive star that collapses when it has used up all its nuclear fuel (types Ib, Ic, and II) or in a white dwarf star that gains enough mass from a nearby companion to exceed the mass limit for electron degeneracy (type Ia).

Supernova remnant: The expanding debris cloud from a star that exploded as a supernova.

Swing-by: The close approach of a spacecraft as it passes by a planet. *See also* Flyby.

Synchronous rotation (or orbit): The state of a body in which the period of rotation equals the average orbital period.

Synodic period: The time between two successive occurrences of the same configuration of three or more bodies.

Système International d'Unités. *See* SI units.

T-Tauri stage: A temporary stage of instability in the early life of a star when it experiences great mass loss and an intense stellar wind.

Taenite: A nickel-iron alloy mineral with more than 25 percent nickel in solid solution; in octahedrite meteorites, taenite forms three-dimensional sheets that appear band-shaped in cut sections.

Tail: The apparent tail of the comet that consists of volatilized, ionized materials carried away from the comet by solar wind and dust blown away from the comet by the radiation pressure of sunlight.

Tectonics: The study of the processes forming the major structural features of a planet's crust.

Tektites: Glassy pellets believed to have formed when rock and soil that were vaporized by extraterrestrial impact recondensed and fell back to Earth.

Telemetry: Real-time transmission of data from a distance via radio signals.

Teleoperations: Manipulation of an orbiter, booster, or instruments in space from Earth via remote control.

Telescope: A device that permits detailed inspection of a distant object; originally, an instrument that used lenses or mirrors to collect large quantities of light and focus it into an image from a tiny area.

Terminator: The line, on a planet or moon, between dark and light that forms the boundary between day and night.

Terrestrial planet: Any of the largely solid, rocky planets of the inner solar system: Mercury, Venus, Earth, and Mars. Sometimes Earth's Moon is included in this list even though it is not strictly a planet.

Tharsis Dome: An immense bulge in the Martian crust in the Tharsis region of the planet, which rises 11 kilometers (6.82 miles) above the Martian surface.

Thermal erosion: The erosion of water ice from a solid state to vapor.

Thermal mapping: Gathering data from which to construct maps by means of instruments capable of sensing heat-producing electromagnetic radiation.

Thermokarst: Geological formations, resembling surface depressions, caused by the melting of subsurface ice or permafrost.

Thermonuclear reaction: A reaction in which atoms of a particular element are fused under pressure and temperature to form atoms of another element, with an explosive release of energy.

Thermosphere: The highest layer of Earth's atmosphere except for the exosphere, beginning at 85 kilometers (52.7 miles) above sea level. The oxygen and nitrogen that compose the atmosphere at this level are extremely rarefied, and are heated by the Sun's ultraviolet radiation to the point of ionization; hence the ionosphere (which lies between 50 and 500 kilometers, or 31 and 310 miles, above sea level) roughly coincides with the thermosphere. This is also the region in which auroras and meteors occur.

Tidal force: The gravitational attraction exerted between two bodies that may potentially distort their shape and cause internal deformation and heating.

Time dilation: The phenomenon, predicted by Albert Einstein's special theory of relativity, whereby time appears to slow down in a system moving near the speed of light from the vantage point of an observer outside that system.

Topside observation: Electronic scanning of Earth's (or another planet's) atmosphere from above. Used mainly in reference to meteorological satellites.

Total lunar eclipse: An eclipse in which the Moon passes completely into the umbra of the Earth.

Total solar eclipse: An eclipse in which the Sun, when viewed from some parts of Earth, is entirely hidden by the Moon.

Transfer orbit: The orbit into which a spacecraft is boosted from Earth orbit on its way to orbit around another celestial body. The spacecraft does not complete a full revolution of the transfer orbit, but only part of the ellipse.

Transit: The alignment of three celestial bodies, especially when a body of small apparent size passes across the disk of larger apparent size as viewed from the third body, such as the rare transit of Mercury or Venus across the Sun as viewed from Earth.

Transponder: A device that receives radio signals and automatically responds to them using the same frequency.

Triangulation: A means of determining the position of an object by calculation from known quantities: the distance between two fixed points and the angles formed between the line described by those points and the line from each of those to a third point, whose distance is not known. Triangulation is the oldest method of determining distances, both on Earth and in space.

Troilite: An iron sulfide mineral that is common to most meteorites but is very rare on Earth.

Troposphere: The lowest level of an atmosphere; it contains the highest density of material in the atmosphere and displays turbulent winds and chemical mixing. Earth's troposphere extends upward from the surface to between 12 and 15 kilometers (7.44 and 9.3 miles); the troposphere contains about 85 percent of the total mass of the atmosphere, almost all the water vapor in the atmosphere, and most of the greenhouse gases.

Tunneling: A quantum mechanical effect in which particles with small but finite probability can pass through a barrier, impenetrable according to Newtonian or "classical" mechanics, and emerge on the other side without energy loss.

Ultraviolet astronomy: The branch of astronomy that examines the ultraviolet emissions of celestial phenomena. Ultraviolet astronomy has developed with the advent of the space age; since ultraviolet rays are unable to penetrate Earth's atmosphere, instruments must be carried aloft by satellites such as the International Ultraviolet Explorer and FUSE.

Ultraviolet radiation: Electromagnetic radiation between about 90 and 400 nanometers which is between X radiation and visible violet light.

Ultraviolet spectrometer: An instrument that measures electromagnetic wavelengths in the ultraviolet range, used on satellites such as the International Ultraviolet Explorer and the Extreme Ultraviolet Explorer.

Umbra: A region of complete darkness, where all parts of the source of illumination are obscured. For example, the darkest inner region in a sunspot or the darkest inner part of the Moon's shadow.

Uniformitarianism: The principle that processes currently operating in nature have always been operating; it suggests that the large-scale features of the Earth were developed very slowly over vast periods of time.

Vacuum: A region containing absolutely nothing. Although a perfect vacuum does not exist in nature, near-vacuum conditions exist in interplanetary, interstellar, and intergalactic space.

Van Allen radiation belts: The two layers of Earth's magnetosphere, discovered by James Van Allen in the late 1950's, in which ionized particles spiral back and forth between Earth's magnetic poles. These zones pose hazards to electronic instruments aboard spacecraft.

Variable star: A star whose brightness varies over time, as a result of several intrinsic or extrinsic factors. There are thousands of such stars, and they can be categorized into seven classes based on the causes of their variation. *See also* Cepheid variable.

Velocity of light, *c*: the speed ($3.00\ 10^8$ meters/second or 186,000 miles per second) of electromagnetic waves, including light, through an empty vacuum.

Vernal equinox: The point on the celestial sphere where the Sun crosses the celestial equator passing from south to north.

Vernier rocket: A small thruster rocket used in space to make fine corrections to a spacecraft's orientation or trajectory.

Very long baseline interferometry: A technique used by radio astronomers to increase perceived sharpness (resolution) of radio sources by electronically linking several radio telescopes at widely spaced locations, allowing accurate mapping of celestial objects that emit radio waves. *See also* Interferometry.

Vesicular: Containing bubble-like cavities.

Virial theorem: The principle, derived from statistical mechanics, that in a cloud contracting gravitationally, the magnitude of the gravitational potential energy must be greater than or equal to twice the total of the thermal, rotational, and magnetic energies. Also known as Jeans's theorem (for Sir James Jeans, one of its developers).

Viscosity: A measure of the ability to flow; a substance with low viscosity flows easily.

Visible spectrum: That portion of the electromagnetic spectrum to which human eyes are sensitive, between about 400 and 700 nanometers, or 0.4 to 0.7 microns.

Visual binaries: Two stars that are gravitationally bound to each other and are either far enough apart or close enough to Earth to enable the two stars to be seen individually through a telescope.

Void: The region of space hundreds of millions of light-years across, which is relatively free of galaxies.

Volatile: Having a low condensation temperature, evaporating readily; the opposite of refractory. Also used in the plural, *volatiles*, as a noun to refer to substances with these properties.

Volcanism: The dynamic process in which molten material from the interior of a planet is transferred to the planet's solid surface, issuing forth explosively from cracks or other openings.

Water of hydration: Molecular water that is bound in the crystalline structure of rocks.

Wavelength: The distance from one extremum displacement to the next one of the same phase in any sort of wave phenomenon, such as water waves, sound waves, or electromagnetic radiation.

White dwarf star: A star with a mass 1.4 times the mass of the Sun or less and a diameter approximately that of Earth, which has exhausted all its nuclear fuels and shines only because it is still hot. As it radiates its energy away, it cools and fades, becoming a black dwarf star.

White spots: Atmospheric disturbances on Jovian planets, appearing at boundaries between zones and belts.

Widmanstätten pattern: A textural pattern, common to most iron meteorites containing 6-16 percent nickel, that occurs when kamacite is oriented parallel to the octahedral planes in taenite.

Wind tunnel: A large, tubular structure through which air is forced to flow at high speeds for the purpose of testing the behavior of aircraft and other structures that travel through the atmosphere.

Work: The product of a force on an object and the component of distance traveled that is parallel to that force.

World line: The graph of the motion of an object through space-time; objects moving with constant velocities have straight world lines, while accelerated motion is represented by curved world lines.

X band: The range of radio frequencies between 5.2 and 10.9 gigahertz.

X radiation: Electromagnetic radiation with wavelengths between 0.1 and 10 nanometers, the range lying between gamma and ultraviolet radiation on the electromagnetic spectrum.

X-ray astronomy: The branch of astronomy that examines the X-ray emissions of celestial phenomena. This radiation cannot be studied from the ground, since Earth's atmosphere absorbs most X radiation, and therefore has blossomed with the advent of spacecraft that can carry X-ray telescopes and other detectors into space. The examination of celestial X-ray sources has led, among other things, to the discovery of neutron stars and black holes as members of binary star systems.

Yaw. *See* Pitch, roll, and yaw.

Zeeman effect: The broadening or splitting of spectral lines of light emitted by atoms caused by the presence of magnetic fields.

Zero gravity: The condition of absolute weightlessness, which occurs in free fall and in orbit. *See also* Microgravity.

Zircons: Mineral inclusions found in granitic rocks, zircons are often the only evidence left of early crustal rocks.

Zodiacal light: The glow seen in the west after sunset and in the east before dawn, caused by sunlight reflecting off microscopic dust particles.

Zones: Atmospheric high pressure bands of yellow-white clouds that encircle Jupiter.

General Bibliography

Abell, George O., David Morrison, and Sidney C. Wolff. *Exploration of the Universe*. 7th ed. Philadelphia: Saunders College Publishing, 1995.

Adams, Fred, and Greg Laughlin. *The Five Ages of the Universe*. New York: Free Press 1999.

Ahrens, C. Donald. *Essentials of Meteorology: An Invitation to the Atmosphere*. 5th ed. Florence, Ky.: Brooks/Cole, 2007.

_______. *Meteorology Today: An Introduction to Weather, Climate, and the Environment*. 8th ed. Florence, Ky.: Brooks/Cole, 2006.

Akasofu, Syun-Ichi, ed. *Dynamics of the Magnetosphere*. Dordrecht, Netherlands: D. Reidel, 1980.

Akasofu, Syun-Ichi, and Y. Kamide, eds. *The Solar Wind and the Earth*. Boston: D. Reidel, 1987.

Alexander, Arthur Francis O'Donel. *The Planet Saturn: A History of Observation, Theory, and Discovery*. New York: Macmillan, 1962.

Allen, Oliver E. *Atmosphere*. Alexandria, Va.: Time-Life Books, 1983.

American Geophysical Union. *Scientific Results of the Viking Project*. Washington, D.C.: Author, 1978.

Anderson, Geoff. *The Telescope: Its History, Technology, and Future*. Princeton, N.J.: Princeton University Press, 2007.

Andrews, David. *An Introduction to Atmospheric Physics*. Cambridge, England: Cambridge University Press, 2000.

Arnett, David. *Supernovae and Nucleosynthesis*. Princeton, N.J.: Princeton University Press, 1996.

Arny, Thomas T., and Stephen E. Schneider. *Explorations: An Introduction to Astronomy*. 5th ed. New York: McGraw-Hill, 2007.

Asimov, Isaac. *The Exploding Suns: The Secrets of the Supernovas*. New York: E. P. Dutton, 1985.

_______. *Understanding Physics*. 1966. Reprint. New York: Barnes & Noble Books, 1993.

Asimov, Isaac, and Richard Hantula. *Jupiter*. Milwaukee, Wis.: Gareth Stevens, 2002.

Asrar, Ghassem. *EOS: Science Strategy for the Earth Observing System*. New York: American Institute of Physics, 1994.

Atreya, S. K., J. B. Pollack, and M. S. Matthus, eds. *Origins and Evolution of Planetary and Satellite Atmospheres*. Phoenix: University of Arizona Press, 1989.

Aveni, Anthony. *Stairways to the Stars: Skywatching in Three Great Ancient Cultures*. New York: Wiley, 1999.

Backus, George. *Foundations of Geomagnetism*. New York: Cambridge University Press, 1996.

Bagenal, Fran, Timothy E. Dowling, and William B. McKinnon, eds. *Jupiter: The Planet, Satellites, and Magnetosphere*. New York: Cambridge University Press, 2007.

Baker, Robert H. *Astronomy*. 7th ed. Princeton, N.J.: Van Nostrand, 1959.

Baker, Victor R. *The Channels of Mars*. Austin: University of Texas Press, 1982.

Baldwin, R. A. *The Measure of the Moon*. Chicago: University of Chicago Press, 1963.

Bally, A. W. *Seismic Expression of Structural Styles*. Tulsa, Okla.: American Association of Petroleum Geologists, 1983.

Bally, John, and Bo Reipurth. *The Birth of Stars and Planets*. Cambridge, England: Cambridge University Press, 2006.

Balogh, André, Leonid Ksanfomality, and Rudolf von Steiger. *Mercury*. New York: Springer, 2008.

Balogh, André, Louis J. Lanzerotti, and S. T. Suess, eds. *The Heliosphere Through the Solar Activity Cycle*. New York: Springer, 2007.

Barbour, Julian B., and Herbert Pfister, eds. *Mach's Principle: From Newton's Bucket to Quantum Gravity*. Boston: Birkhäuser, 1995.

Barlow, Nadine. *Mars: An Introduction to Its Interior, Surface, and Atmosphere*. Cambridge, England: Cambridge University Press, 2008.

Barnes-Svarney, Patricia. *Asteroid: Earth Destroyer or New Frontier?* New York: Basic Books, 2003.

Barrow, John D., and Joseph Silk. *The Left Hand of Creation*. New York: Basic Books, 1983.

Barstow, Martin A., and Jay B. Holberg. *Extreme Ultraviolet Astrophysics*. Cambridge, England: Cambridge University Press, 2003.

Bartusiak, Marcia. *Thursday's Universe*. New York: Times Books, 1986.

Bascom, Willard. *A Hole in the Bottom of the Sea*. Garden City, N.Y.: Doubleday, 1961.

Beattie, Donald A. *Taking Science to the Moon: Lunar Experiments and the Apollo Program*. Baltimore: Johns Hopkins University Press, 2003.

Beatty, J. Kelly, Carolyn Collins Petersen, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge, Mass.: Sky, 1999.

Beiser, Arthur. *Concepts of Modern Physics*. 6th ed. New York: McGraw-Hill, 2002.

Bell, Jim. *Postcards from Mars: The First Photographer on the Red Planet*. New York: Dutton Adult, 2006.

_______, ed. *The Martian Surface*. Cambridge, England: Cambridge University Press, 2008.

Bell, Jim, and Jacqueline Mitton, eds. *Asteroid Rendezvous: NEAR Shoemaker's Adventures at Eros*. Cambridge, England: Cambridge University Press, 2002.

Belusevic, Radoje. *Relativity, Astrophysics, and Cosmology*. Weinheim, Germany: Wiley-VCH, 2008.

Bely, Pierre, ed. *The Design and Construction of Large Optical Telescopes*. New York: Springer, 2003.

Bennett, Jeffrey, et al. *The Cosmic Perspective*. 5th ed. San Francisco: Pearson/Addison-Wesley, 2008.

_______. *Stars, Galaxies, and Cosmology: The Cosmic Perspective*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2004.

Benton, Julius. *Saturn and How to Observe It*. New York: Springer, 2005.

Bernal, J. D. *The Origin of Life*. New York: World Books, 1967.

Bertin, G., and C. C. Lin. *Spiral Structure in Galaxies: A Density Wave Theory*. Cambridge, Mass.: MIT Press, 1996.

Beskin, V. S., A. V. Gurevich, and Ya N. Istomin. *Physics of the Pulsar Magnetosphere*. Cambridge, England: Cambridge University Press, 2006.

Beutler, G., M. R. Drinkwater, R. Rummel, and Rudolf von Steiger. *Earth Gravity Field from Space: From Sensors to Earth Sciences*. Boston: Kluwer Academic, 2003.

Bevan, Alex, and John De Laeter. *Meteorites: A Journey Through Space and Time*. Washington, D.C.: Smithsonian Institution Press, 2002.

_______. *Thunderstones and Shooting Stars: The Meaning of Meteorites*. Cambridge, Mass.: Harvard University Press, 1986.

Bhatnagar, Arvind, and William Livingston. *Fundamentals of Solar Astronomy*. Hackensack, N.J.: World Scientific, 2005.

Blondel, Phillippe, and John Mason, eds. *Solar System Update*. New York: Springer, 2006.

Bobrowsky, Peter T., and Hans Rickman, eds. *Comet/Asteroid Impacts and Human Society: An Interdisplinary Approach*. New York: Springer, 2007.

Bok, Bart J., and Priscilla F. Bok. *The Milky Way*. 5th ed. Cambridge, Mass.: Harvard University Press, 1981.

Bolt, Bruce A. *Earthquakes*. New York: W. H. Freeman, 1988.

_______. *Inside the Earth: Evidence from Earthquakes*. San Francisco: W. H. Freeman, 1982.

Bond, Peter. *Distant Worlds: Milestones in Planetary Exploration*. New York: Copernicus Books, 2007.

Bone, Neil. *The Aurora: Sun-Earth Interaction*. New York: John Wiley, 1996.

Bortolotti, Dan. *Exploring Saturn*. New York: Firefly Books, 2003.

Bostrom, Robert C. *Tectonic Consequences of the Earth's Rotation*. New York: Oxford University Press, 2000.

Bothmer, Volker, and Ioannis A. Daglis. *Space Weather: Physics and Effects*. New York: Springer Praxis, 2006.

Bott, M. H. P. *The Interior of the Earth*. New York: Elsevier, 1982.

Bottke, William F., Jr., Albertoi Cellino, Paolo Paolicchi, and Richard P. Binzel, eds. *Asteroids III*. Tucson: University of Arizona Press, 2002.

Boucher, C., ed. *Earth Rotation and Coordinate Reference Frames*. International Association of Geodesy Symposia 105. New York: Springer, 1990.

Boyce, Joseph M. *The Smithsonian Book of Mars*. Washington, D.C.: Smithsonian Institution Press, 2002.

Brancazio, Peter J., ed. *The Origin and Evolution of Atmospheres and Oceans*. New York: John Wiley & Sons, 1964.

Brandt, John C., and Robert D. Chapman. *Introduction to Comets*. New York: Cambridge University Press, 2004.

Bredeson, Carmen. *NASA Planetary Spacecraft: Galileo, Magellan, Pathfinder, and Voyager*. New York: Enslow, 2000.

Brewer, Bryan. *Eclipse*. 2d ed. Seattle: Earth View, 1991.

Briggs, G. A., and F. W. Taylor. *The Cambridge Photographic Atlas of the Planets*. Cambridge, England: Cambridge University Press, 1982.

Brody, Judit. *The Enigma of Sunspots: A Story of Discovery and Scientific Revolution*. Edinburgh, Scotland: FlorisSunspots, 2002.

Brown, G. C. *The Inaccessible Earth: An Integrated View to Its Structure and Composition*. 2d ed. New York: Chapman and Hall, 1993.

Brown, Peter L. *Comets, Meteorites, and Men*. New York: Taplinger, 1975.

Brunier, Serge, and Anne-Marie Lagrange. *Great Observatories of the World*. Buffalo, N.Y.: Firefly Books, 2005.

Brusche, P., and Sundermann, J. *Tidal Friction and the Earth's Rotation*. Berlin: Springer, 1978.

Brush, Stephen G. *Nebulous Earth: The Origin of the Solar System and the Core of the Earth from Laplace to Jeffreys*. Cambridge, England: Cambridge University Press, 1996.

Buchler, J. Robert, and Henry Kandrop, eds. *Astrophysical Turbulence and Convection*. New York: New York Academy of Sciences, 2000.

Buchwald, Vagn F. *Handbook of Iron Meteorites: Their History, Distribution, Composition, and Structure*. Berkeley: University of California Press, 1975.

Burgess, Eric. *Uranus and Neptune: The Distant Giants*. New York: Columbia University Press, 1988.

_______. *Venus: An Errant Twin*. New York: Columbia University Press, 1985.

Burke, Bernard F., and Francis Graham-Smith. *An Introduction to Radio Astronomy*. 2d ed. Cambridge, England: Cambridge University Press, 2002.

Burke, John G. *Cosmic Debris: Meteorites in History*. Berkeley: University of California Press, 1986.

Cairns-Smith, A. G. *Genetic Takeover and the Mineral Origins of Life*. Cambridge, England: Cambridge University Press, 1982.

Calder, Nigel. *Einstein's Universe*. New York: Viking Press, 1979.

Cameron, A. G. W., ed. *Interstellar Communication*. New York: W. A. Benjamin, 1963.

Canup, R. M., and K. Righter, eds. *Origin of the Earth and Moon*. Tucson: University of Arizona Press, 2000.

Carr, Michael H. *The Surface of Mars*. Cambridge, England: Cambridge University Press, 2006.

Carroll, Bradley W., and Dale A. Ostlie. *An Introduction to Modern Astrophysics*. 2d ed. San Francisco: Pearson/Addison-Wesley, 2007.

Casoli, Fabienne, and Thérèse Encrenaz. *The New Worlds: Extrasolar Planets*. New York: Springer Praxis, 2007.

Cattermole, Peter John. *Venus: The Geological Story*. Baltimore: Johns Hopkins University Press, 1996.

Chaikin, Andrew. *A Man on the Moon: The Voyages of the Apollo Astronauts*. New York: Penguin, 2007.

Chaisson, Eric. *Cosmic Dawn: The Origins of Matter and Life*. Boston: Little, Brown, 1981.

_______. *The Hubble Wars*. New York: HarperCollins, 1994.

_______. *Relatively Speaking*. New York: W. W. Norton, 1988.

Chaisson, Eric, and Steve McMillan. *Astronomy: A Beginner's Guide to the Universe*. 5th ed. Upper Saddle River, N.J.: Pearson/Prentice Hall, 2007.

_______. *Astronomy Today*. 6th ed. New York: Addison-Wesley, 2008.

Chamberlain, Joseph W. *Theory of Planetary*

Atmospheres: An Introduction to Their Physics and Chemistry. New York: Academic Press, 1978.

Chamberlain, Von Dei, John Carlson, and M. Jane Young. *Songs from the Sky: Indigenous Astronomical and Cosmological Traditions of the World*. West Sussex, England: Ocarina Books, 2005.

Christensen, Lars Lindberg, Robert A. Fosbury, and M. Kornmesser. *Hubble: Fifteen Years of Discovery*. New York: Springer, 2006.

Christian, James L. *Extra-terrestrial Intelligence: The First Encounter*. Buffalo, N.Y.: Prometheus, 1976.

Clark, Pamela. *Dynamic Planet: Mercury in the Context of Its Environment*. New York: Springer, 2007.

Clayton, Donald. *Handbook of Isotopes in the Cosmos: Hydrogen to Gallium*. Cambridge, England: Cambridge University Press, 2003.

Close, Frank. *Apocalypse When? Cosmic Catastrophe and the Fate of the Universe*. New York: William Morrow, 1988.

Cohen, Jack, and Ian Stewart. *Evolving the Alien: The Science of Extraterrestrial Life*. London: Ebury Press, 2002.

Cohen, Martin. *In Darkness Born: The Study of Star Formation*. New York: Cambridge University Press, 1988.

Cohen, Nathan. *Gravity's Lens: Views of the New Cosmology*. New York: John Wiley & Sons, 1988.

Cole, George H. A., and Michael M. Woolfson, eds. *Planetary Science: The Science of Planets Around Stars*. Bristol: Institute of Physics, 2002.

Cole, Michael D. *Galileo Spacecraft: Mission to Jupiter*. New York: Enslow, 1999.

Collins, Michael. *Mission to Mars*. New York: Grove Weidenfeld, 1990.

Comins, N. *What if the Moon Didn't Exist? Voyages to Earths That Might Have Been*. New York: HarperCollins, 1993.

Condie, Kent, and Robert Sloan. *Origin and Evolution of Earth: Principles of Historical Geology*. Upper Saddle River, N.J.: Prentice Hall, 1998.

Consolmagno, Guy, and Martha Schaefer. *Worlds Apart: A Textbook in Planetary Sciences*. Englewood Cliffs, N.J.: Prentice Hall, 1994.

Cook, Alex. *The Greenhouse Effect: A Legacy*. Indianapolis: Dog Ear, 2007.

Cooke, Donald A. *The Life and Death of Stars*. New York: Crown, 1985.

Corfield, Richard. *Lives of the Planets: A Natural History of the Solar System*. New York: Basic Books, 2007.

Corliss, William R., ed. *The Moon and the Planets*. Glen Arm, Md.: Sourcebook Project, 1985.

Cornell, James, and Alan P. Lightman. *Revealing the Universe*. Cambridge, Mass.: MIT Press, 1982.

Coustenis, Athena, and Fredric W. Taylor. *Titan: Exploring an Earthlike World*. Hackensack, N.J.: World Scientific, 2007.

Cox, A. N., W. C. Livingston, and M. S. Matthews, eds. *Solar Interior and Atmosphere*. Tucson: University of Arizona Press, 1991.

Cramer, John A. *How Alien Would Aliens Be?* Lincoln, Nebr.: Writers Club Press, 2001.

Crick, Francis. *Life Itself: Its Origin and Nature*. New York: Simon & Schuster, 1981.

Cross, Charles A., and Patrick Moore. *The Atlas of Mercury*. New York: Crown, 1977.

Croswell, Ken. *Ten Worlds: Everything That Orbits the Sun*. Honesdale, Pa.: Boyds Mills Press, 2007.

Cruikshank, Dale P., ed. *Neptune and Triton*. Tucson: University of Arizona Press, 1995.

Dalrymple, G. Brent. *Ancient Earth, Ancient Skies: The Age of the Earth and Its Cosmic Surroundings*. Stanford, Calif.: Stanford University Press, 2004.

Darwin, G. *The Tides and Kindred Phenomena in the Solar System*. San Francisco: W. H. Freeman, 1962.

Dasch, Pat, ed. *Icy Worlds of the Solar System*. Cambridge, England: Cambridge University Press, 2004.

Dauber, Philip M., and Richard A. Muller. *The Three Big Bangs*. Reading, Mass.: Addison-Wesley, 1996.

Davies, Ashley Gerard. *Volcanism on Io: A Companion with Earth*. Cambridge, England: Cambridge University Press, 2007.

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

Davies, Merton E., Stephen E. Dwornik, Donald E. Gault, and Robert G. Strom. *Atlas of Mercury*. NASA SP-423. Washington, D.C.: National Aeronautics and Space Administration, Scientific and Technical Information Office, 1978.

Davies, P. C. W. *Space and Time in the Modern Universe*. New York: Cambridge University Press, 1981.

De Bremaecker, Jean-Claude. *Geophysics: The Earth's Interior*. New York: John Wiley & Sons, 1985.

Deeg, Hans, Juan Antonio Belmonte, and Antonio Aparicio, eds. *Extrasolar Planets*. New York: Cambridge University Press, 2008.

Delobeau, Francis. *The Environment of the Earth*. New York: Springer, 1971.

Delsemme, A. H., ed. *Comets, Asteroids, Meteorites*. Toledo, Ohio: University of Toledo Press, 1977.

De Pater, Imke, and Jack J. Lissauer. *Planetary Sciences*. New York: Cambridge University Press, 2001.

Dermott, S. F., ed. *The Origin of the Solar System*. New York: John Wiley & Sons, 1978.

Devorkin, David, and Robert W. Smith. *Hubble: Imaging Space and Time*. Washington, D.C.: National Geographic Society, 2008.

Dick, Steven J. *The Biological Universe: The Twentieth-Century Extraterrestrial Life Debate and the Limits of Science*. Cambridge, England: Cambridge University Press, 1996.

Dinwiddie, Robert, et al. *Universe*. New York: DK Adult, 2005.

Dixon, Dougal. *The Practical Geologist: The Introductory Guide to the Basics of Geology and to Collecting and Identifying Rocks*. New York: Fireside, 1992.

Dodd, Robert T. *Meteorites: A Petrologic-Chemical Synthesis*. London: Cambridge University Press, 1981.

_______. *Thunderstones and Shooting Stars: The Meaning of Meteorites*. Cambridge, Mass.: Harvard University Press, 1986.

Dole, Stephen H. *Habitable Planets for Man*. 2d ed. New York: Elsevier, 1970.

Domingue, D. L., and C. T. Russell, eds. *The MESSENGER Mission to Mercury*. New York: Springer, 2008.

Drexler, Jerome. *Discovering Postmodern Cosmology: Discoveries in Dark Matter, Cosmic Web, Big Bang, Inflation, Cosmic Rays, Dark Energy, Accelerating Cosmos*. Boca Raton, Fla.: Universal, 2008.

Dudly, W. W., and D. A. Williams. *Interstellar Chemistry*. New York: Academic Press, 1984.

Duncan, Todd, and Craig Tyler. *Your Cosmic Context: An Introduction to Modern Cosmology*. San Francisco: Pearson/Addison-Wesley, 2009.

Dunne, James A., and Eric Burgess. *The Voyage of Mariner 10: Mission to Venus and Mercury*. NASA SP-424. Washington, D.C.: National Aeronautics and Space Administration, Scientific and Technical Information Office, 1978.

Dvorak, Rudolf. *Extrasolar Planets: Formation, Detection, and Dynamics*. Weinheim, Germany: Wiley-VCH, 2008.

Eddington, Arthur. *The Expanding Universe*. 1933. Reprint. Cambridge, England: Cambridge University Press, 1987.

Eddy, John A. *A New Sun: The Solar Results from Skylab*. Washington, D.C.: Government Printing Office, 1979.

Elkins-Tanton, Linda T. *Jupiter and Saturn*. New York: Chelsea House, 2006.

_______. *Mars*. New York: Chelsea House, 2006.

_______. *The Sun, Mercury, and Venus*. New York: Chelsea House, 2006.

_______. *Uranus, Neptune, Pluto, and the Outer Solar System*. New York: Chelsea House, 2006.

Elliot, James, and Richard Kerr. *Rings: Discoveries from Galileo to Voyager*. Cambridge, Mass.: MIT Press, 1984.

Ellison, Mervyn Archdall. *The Sun and Its Influence*. London: Routledge & Kegan Paul, 1955.

Emiliani, Cesare. *The Scientific Companion*. New York: John Wiley and Sons, 1988.

Encrenaz, Thérèse, Jean-Pierre Bibring, Michel Blanc, and Maria-Antonietta Barucci. *The Solar System*. New York: Springer, 2004.

Encrenaz, Thérèse, Reinald Kallenbach, T. Owen, and C. Sotin, eds. *The Outer Planets and Their Moons: Comparative Studies of the Outer Planets Prior to the Exploration of the Saturn System by Cassini-Huygens*. New York: Springer, 2005.

Erickson, Jon. *Asteroids, Comets, and Meteorites: Cosmic Invaders of the Earth*. New York: Facts On File, 2003.

Esposito, Larry. *Planetary Rings*. Cambridge, England: Cambridge University Press, 2006.

Esposito, Larry W., Ellen R. Stofan, and Thomas E. Cravens, eds. *Exploring Venus as a Terrestrial Planet: Geophysical Monograph 176*. Washington, D.C.: American Geophysical Union, 2007.

European Space Agency. *The Atmospheres of Saturn and Titan*. ESA SP-241. Paris: Author, 1985.

Ezell, Edward, and Linda Ezell. *On Mars: Explorations of the Red Planet, 1958-1978*. NASA SP-4212. Washington, D.C.: Government Printing Office, 1984.

Fabian, A. C., K. A. Pounds, and R. D. Blandford. *Frontiers of X-ray Astronomy*. Cambridge, England: Cambridge University Press, 2004.

Fairbridge, Rhodes W. *The Encyclopedia of Geochemistry and Environmental Sciences*. Stroudsburg, Pa.: Bowden, Hutchinson and Ross, 1972.

Faure, Gunter, and Teresa M. Mensing. *Introduction to Planetary Science: The Geological Perspective*. New York: Springer, 2007.

Ferguson, Kitty. *Tycho and Kepler: The Unlikely Partnership That Forever Changed Our Understanding of the Heavens*. New York: Walker, 2002.

Ferington, Esther. *The Cosmos*. Alexandria, Va.: Time-Life Books, 1988.

Fernández, Julio Angel. *Comets: Nature, Dynamics, Origin, and Their Cosmological Relevance*. Dordrecht, Netherlands: Springer, 2005.

Ferreira, Pedro G. *The State of the Universe: A Primer in Modern Cosmology*. London: Phoenix, 2007.

Field, George. *The Space Telescope*. Chicago: Contemporary Books, 1989.

Fimmel, Richard O., Lawrence Colin, and Eric Burgess. *Pioneering Venus: A Planet Unveiled*. Washington, D.C.: National Aeronautics and Space Administration, 1995.

Fimmel, Richard O., James Van Allen, and Eric Burgess. *Pioneer: First to Jupiter, Saturn and Beyond*. NASA SP-446. Washington, D.C.: Government Printing Office, 1980.

Fischer, Daniel. *Mission Jupiter: The Spectacular Journey of the Galileo Spacecraft*. New York: Copernicus Books, 2001.

Foerstner, Abigail. *James Van Allen: The First Eight Billion Miles*. Iowa City: University of Iowa Press, 2007.

Forget, Françoise, Françoise Costard, and Philippe Lognonné. *Planet Mars: Story of Another World*. Chichester, England: Praxis, 2008.

Foster, J., and J. D. Nightingale. *A Short Course in General Relativity*. 2d ed. New York: Springer, 1995.

Foukal, Peter. *Solar Astrophysics*. 2d rev. ed. Weinheim, Germany: Wiley-VCH, 2004.

Fountain, John, and Rolf Sinclair. *Current Studies in Archaeoastronomy: Conversations Across Time and Space*. Durham, N.C.: Carolina Academic Press, 2005.

Fowler, C. M. R. *The Solid Earth: An Introduction to Global Geophysics*. 2d ed. New York: Cambridge University Press, 2004.

Fowler, G. C. *The Inaccessible Earth: An Integrated View to Its Structure and Composition*. 2d ed. New York: Chapman and Hall, 1993.

Fowles, Grant R., and George L. Cassiday. *Analytic Mechanics*. 7th ed. New York: Brooks/Cole, 2004.

Frakes, L. A. *Climates Throughout Geologic Time*. New York: Elsevier, 1980.

Fraknoi, Andrew, David Morrison, and Sidney Wolff. *Voyages to the Stars and Galaxies*. Belmont, Calif.: Brooks/Cole-Thomson Learning, 2006.

Frazier, Kendrick. *Our Turbulent Sun*. Englewood Cliffs, N.J.: Prentice-Hall, 1982.

_______. *Solar Systems*. Alexandria, Va.: Time-Life Books, 1985.

Freedman, Roger A., and William J. Kaufmann III. *Universe*. 8th ed. New York: W. H. Freeman, 2007.

French, Bevan M. *The Moon Book*. New York: Penguin Books, 1977.

Fridman, Alexei M., and Nikolai N. Gorkavyi. *Physics of Planetary Rings: Celestial Mechanics of Continuous Media*. New York: Springer, 1999.

Friedlander, Michael W. *Cosmic Rays*. Cambridge, Mass.: Harvard University Press, 1989.

Friedman, Herbert. *Sun and Earth*. San Francisco: W. H. Freeman, 1986.

Fukugita, Masataka, and Tsutomu Yanagida. *Physics of Neutrinos*. New York: Springer, 2003.

Fyfe, W. S. *Geochemistry*. Oxford, England: Clarendon Press, 1974.

Gabler, Robert E., Robert J. Sager, Sheila M. Brazier, and D. L. Wise. *Essentials of Physical Geography*. 8th ed. Florence, Ky.: Brooks/Cole, 2006.

Gamow, George. *Gravity*. New York: Dover, 2003.

Garland, G. D. *Introduction to Geophysics*. 2d ed. Philadelphia: W. B. Saunders, 1979.

Garlick, Mark A. *The Story of the Solar System*. Cambridge, England: Cambridge University Press, 2002.

Garrison, Tom. *Oceanography: An Invitation To Marine Science*. Florence, Ky.: Brooks/Cole, 2007.

Gasperini, Maurizio. *The Universe Before the Big Bang: Cosmology and String Theory*. Berlin: Springer, 2008.

Gehrels, Tom, ed. *Asteroids*. Tucson: University of Arizona Press, 1979.

_______. *Hazards Due to Comets and Asteroids*. Tucson: University of Arizona Press, 1994.

_______. *Jupiter*. Tucson: University of Arizona Press, 1976.

Gehrels, Tom, and Mildred Shapley Matthews, eds. *Saturn*. Tucson: University of Arizona Press, 1984.

Genet, Russell M., Donald S. Hayes, Douglas S. Hall, and David R. Genet. *Supernova 1987a: Astronomy's Explosive Enigma*. Mesa, Ariz.: Fairborn Press, 1988.

Gibson, Edward G. *The Quiet Sun*. NASA SP-303. Washington, D.C.: Government Printing Office, 1973.

Ginzburg, V. L., and S. I. Syrovatskii. *The Origin of Cosmic Rays*. New York: Macmillan, 1964.

Giovanelli, Ronald G. *Secrets of the Sun*. New York: Cambridge University Press, 1984.

Giunti, Carlo, and Chung W. Kim. *Fundamentals of Neutrino Physics and Astrophysics*. New York: Oxford University Press, 2007.

Glasstone, Samuel. *The Book of Mars*. Washington, D.C.: National Aeronautics and Space Administration, 1968.

Glendenning, Norman K. *Compact Stars: Nuclear Physics, Particle Physics, and General Relativity*. New York: Springer, 1997.

Goldsmith, Donald. *Worlds Unnumbered: The Search for Extrasolar Planets*. Sausalito, Calif.: University Science Books, 1997.

Goldsmith, Donald, and Tobias Owen. *The Search for Life in the Universe*. 3d ed. New York: University Science Books, 2001.

Golub, Leon, and Jay M. Pasachoff. *Nearest Star: The Surprising Science of Our Sun*. Cambridge, Mass.: Harvard University Press, 2001.

Gonzalez, Guillermo, and Jay W. Richards. *The Privileged Planet: How Our Place in the Cosmos Is Designed for Discovery*. Washington, D.C.: Regnery, 2004.

Goody, R. M., and J. C. G. Walker. *Atmospheres*. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Gould, Stephen Jay. *The Flamingo's Smile*. New York: W. W. Norton, 1985.

Greely, R. *Planetary Landscapes*. 2d ed. Boston: Allen and Unwin, 1994.

Green, Simon F., Mark H. Jones, and S. Jocelyn Burnell. *An Introduction to the Sun and Stars*. New York: Cambridge University Press, 2004.

Greenberg, John L. *The Problem of the Earth's Shape from Newton to Clairaut*. New York: Cambridge University Press, 1995.

Greenberg, Richard. *Europa the Ocean Moon: Search for an Alien Biosphere*. New York: Springer, 2005.

Greene, Brian. *The Elegant Universe: Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory*. New York: W. W. Norton, 2003.

Greenstein, George. *Frozen Star*. New York: Charles Scribner's Sons, 1983.

Grego, Peter. *The Moon and How to Observe It*. New York: Kindle Books, 2005.

Gregor, C. Bryan, et al. *Chemical Cycles in the Evolution of the Earth*. New York: John Wiley & Sons, 1988.

Gregory, Stephen A. *Introductory Astronomy and Astrophysics*. 4th ed. San Francisco: Brooks/Cole, 1997.

Grewing, M., F. Praderie, and R. Reinhard, eds. *Exploration of Halley's Comet*. New York: Springer, 1989.

Gribbin, John. *Blinded by the Light: The Secret Life of the Sun*. New York: Harmony Books, 1991.

_______. *In Search of the Big Bang*. New York: Bantam Books, 1986.

_______. *Spacewarps*. New York: Delacorte Press, 1984.

_______. *Timewarps*. New York: Delacorte Press, 1980.

Gribbin, John, and Mary Gribbin. *Fire on Earth*. New York: St. Martin's Press, 1996.

_______. *Stardust: Supernovae and Life, the Cosmic Connection*. New Haven, Conn.: Yale University Press, 2000.

Grinspoon, David. *Lonely Planets: The Natural Philosophy of Alien Life*. New York: HarperCollins, 2004.

_______. *Venus Revealed: A New Look Below the Clouds of Our Mysterious Twin Planet*. New York: Basic Books, 1998.

Gussinov, Oktay, Efe Yazgan, and Askin Ankay, eds. *Neutron Stars, Supernovae, and Supernovae Remnants*. New York: Nova Science, 2007.

Haber, Frances C. *The Age of the World: Moses to Darwin*. 1959. Reprint. Westport, Conn.: Greenwood Press, 1978.

Haigh, Joanna, et al. *The Sun, Solar Analogs, and the Climate*. New York: Springer, 2005.

Halliday, David, Robert Resnick, and Jearl Walker. *Fundamentals of Physics, Extended*. 9th ed. New York: Wiley, 2007.

Hamblin, W. Kenneth, and Eric H. Christiansen. *Exploring the Planets*. New York: Macmillan, 1990.

Hammel, H. B. *The Ice Giant Systems of Uranus and Neptune*. New York: Springer, 2006.

Hansen, Carl J., Steven D. Kawaler, and Virginia Trimble. *Stellar Interiors: Physical Principles, Structures, Evolution*. New York: Springer, 2004.

Hansen, Joel E., and T. Takahashi, eds. *Climate Processes and Climate Sensitivity*. Geophysical Monograph 29. Washington, D.C.: American Geophysical Union, 1984.

Hansson, Anders. *Mars and the Development of Life*. New York: Ellis Horwood, 1991.

Hargreaves, John K. *The Upper Atmosphere and Solar-Terrestrial Relations*. New York: Van Nostrand Reinhold, 1979.

Harland, David M. *Cassini at Saturn: Huygens Results*. New York: Springer, 2007.

_______. *Jupiter Odyssey: The Story of NASA's Gelileo Mission*. New York: Springer, 2000.

_______. *Mission to Saturn: Cassini and the Huygens Probe*. New York: Springer Praxis, 2002.

_______. *Water and the Search for Life on Mars*. New York: Springer Praxis, 2005.

Harpur, Brian, and Laurence Anslow. *The Official Halley's Comet Project Book*. London: Hodder and Stoughton, 1985.

Hartmann, William K. *Moons and Planets*. 5th ed. Belmont, Calif.: Thomson Brooks/Cole, 2005.

_______. *A Traveler's Guide to Mars: The Mysterious Landscapes of the Red Planet*. New York: Workman, 2003.

_______, ed. *Astronomy*. 5th ed. Belmont, Calif.: Wadsworth, 2004.

Hartmann, William K., and Chris Impey. *Astronomy: The Cosmic Journey*. New York: Brooks/Cole, 2001.

Hartmann, William K., and Ron Miller. *The Grand Tour: A Traveler's Guide to the Solar System*. 3d ed. New York: Workman, 2005.

Hartmann, William K., and Odell Raper. *The New Mars: The Discoveries of Mariner 9*. NASA SP-337. Washington, D.C.: Government Printing Office, 1974.

Hartmann, William K., et al. *Out of the Cradle: Exploring the Frontiers Beyond Earth*. New York: Workman, 1984.

Harvey, Brian. *Russian Planetary Exploration: History, Development, Legacy, and Prospects*. New York: Springer, 2007.

Harwit, Martin. *Cosmic Discovery*. New York: Basic Books, 1981.

Hawking, Stephen. *A Brief History of Time*. New York: Bantam Books, 1988.

_______. *An Even Briefer History of Time*. New York: Bantam, 2008.

Hawking, Stephen W., and William Israel.

Three Hundred Years of Gravitation. New York: Cambridge University Press, 1987.
Hawking, Stephen, and Roger Penrose. *The Nature of Space and Time*. Princeton, N.J.: Princeton University Press, 2000.
Hawking, Stephen, et al. *The Future of Spacetime*. New York: W. W. Norton, 2003.
Hazen, Robert. *Genesis: The Scientific Quest for Life's Origins*. Washington, D.C.: Joseph Henry Press, 2005.
Heiken, Grant, and Eric Jones. *On the Moon: The Apollo Journals*. New York: Springer, 2007.
Henbest, Nigel. *Mysteries of the Universe*. New York: Van Nostrand Reinhold, 1981.
Henbest, Nigel, and Michael Marten. *The New Astronomy*. New York: Cambridge University Press, 1983.
Henderson-Sellers, A. *The Origin and Evolution of Planetary Atmospheres*. Bristol, England: Adam Hilger, 1983.
_______, ed. *Satellite Sensing of a Cloudy Atmosphere: Observing the Third Planet*. London: Taylor and Francis, 1984.
Hester, Jeff, George Blumenthal, David Burstein, and Bradford Smith. *Twenty-first Century Astronomy*. New York: W. W. Norton, 2007.
Hey, J. S. *The Evolution of Radio Astronomy*. New York: Science History Publications, 1973.
Hilbrecht, Heinz, Klaus Reinsch, Peter Volker, and Rainer Beck. *Solar Astronomy Handbook*. New York: Willmann-Bell, 1995.
Hille, Steele, and Michael Carlowicz. *The Sun*. New York: Harry N. Abrams, 2006.
Hillier, Rodney. *Gamma-Ray Astronomy*. New York: Oxford University Press, 1984.
Hirsch, Richard F. *Glimpsing an Invisible Universe: The Emergence of X-Ray Astronomy*. New York: Cambridge University Press, 1983.
Hofmann-Wellenhof, Bernhard. *Physical Geodesy*. New York: Springer, 2006.
Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans*. Princeton, N.J.: Princeton University Press, 1984.
Horton, E., and John H. Jones, eds. *Origin of the Earth*. New York: Oxford University Press, 1990.
Hoyt, Douglas V., and Kenneth H Schatten. *The Role of the Sun in Climate Change*. Oxford: Oxford University Press, 1997.
Hoyt, William G. *Lowell and Mars*. Tucson: University of Arizona Press, 1976.
Hubbard, William B. *Planetary Interiors*. New York: Van Nostrand Reinhold, 1984.
Hunt, Garry E., and Patrick Moore. *Atlas of Neptune*. Cambridge, England: Cambridge University Press, 1994.
_______. *Atlas of Saturn*. London: Mitchell Beazley, 1982.
_______. *Atlas of Uranus*. New York: Cambridge University Press, 1988.
_______. *Jupiter*. New York: Rand McNally, 1981.
Hurley, Patrick M. *How Old Is the Earth?* Garden City, N.Y.: Doubleday, 1959.
Hutchison, Robert. *The Search for Our Beginning: An Enquiry, Based on Meteorite Research, into the Origin of Our Planet and Life*. Oxford, England: Oxford University Press, 1983.

Ince, Martin. *The Rough Guide to the Earth 1*. New York Rough Guides, 2007.
Inglis, Mike. *Observer's Guide to Stellar Evolution*. London: Springer, 2003.
Irwin, Patrick G. J. *Giant Planets of Our Solar System: An Introduction*. 2d ed. New York: Springer, 2006.

Jacobs, J. A. *The Earth's Core*. 2d ed. New York: Academic Press, 1987.
Jacobs, John A., Richard D. Russell, and J. T. Wilson. *Physics and Geology*. 2d ed. New York: McGraw-Hill, 1974.
James, David E., ed. *The Encyclopedia of Solid Earth Geophysics*. New York: Van Nostrand Reinhold, 1989.
Jastrow, Robert. *God and the Astronomers*. New York: Warner Books, 1978.
_______. *Red Giants and White Dwarfs*. New York: W. W. Norton, 1979.
Jöels, Kerry Mark. *The Mars One Crew Manual*. New York: Ballantine Books, 1985.
Johnson, Francis S., ed. *Satellite Environment Handbook*. 2d ed. Stanford, Calif.: Stanford University Press, 1965.

Jones, Barrie W. *Discovering the Solar System*. New York: John Wiley & Sons, 1999.

Jordan, Stuart, ed. *The Sun as a Star*. NASA SP-450. Washington, D.C.: Government Printing Office, 1981.

Kaler, James B. *The Hundred Greatest Stars*. New York: Copernicus Books, 2002.

Kallenbach, Reinald, Thérèse Encrenaz, J. Geiss, and Konrad Mauersberger, eds. *Solar System History from Isotopic Signatures of Volatile Elements*. New York: Springer, 2003.

Kallmann-Bijl, Hildegaard, ed. *Space Research: Proceedings of the First International Space Science Symposium*. New York: Interscience, 1960.

Kargel, Jeffrey S. *Mars: A Warmer, Wetter Planet*. New York: Springer Praxis, 2004.

Karttunen, Hannu P., Pekka Kröger, Heikku Oja, and Markku Poutanen, eds. *Fundamental Astronomy*. 5th ed. New York: Springer, 2007.

Katz, Johnathan I. *High Energy Astrophysics*. Reading, Mass.: Addison-Wesley, 1987.

Kaula, William M. *Theory of Satellite Geodesy: Applications of Satellites to Geodesy*. New York: Dover, 2000.

Kelley, David, and Eugene Milone. *Exploring Ancient Skies: An Encyclopedia Survey of Archaeoastronomy*. New York: Springer, 2004.

Kenyon, Ian. *The Light Fantastic: A Modern Introduction to Classical and Quantum Optics*. New York: Oxford University Press, 2008.

Kerridge, John F., and Mildred S. Matthews, eds. *Meteorites and the Early Solar System*. Tucson: University of Arizona Press, 1988.

Kerrod, Robin. *Uranus, Neptune, and Pluto*. New York: Lerner, 2000.

Kerrod, Robin, Carole Stott, and David S. Leckrone. *Hubble: The Mirror on the Universe*. New York: Firefly Books, 2007.

Kieffer, Hugh H., Bruce M. Jakowsky, Conway W. Snyder, and Mildred Matthews, eds. *Mars*. Tucson: University of Arizona Press, 1992.

King, Elbert A., Jr. *Space Geology: An Introduction*. New York: John Wiley and Sons, 1976.

Kippenhahn, Rudolf. *Light from the Depths of Time*. New York: Springer, 1987.

_______. *One Hundred Billion Suns: The Birth, Life, and Death of the Stars*. New York: Basic Books, 1985.

Kirby-Smith, Henry T. *U.S. Observatories: A Directory and Travel Guide*. New York: Van Nostrand Reinhold, 1976.

Kitchin, Christopher R. *Astrophysical Techniques*. 5th ed. Boca Raton, Fl.: CRC Press, 2009.

_______. *Solar Observing Techniques*. New York: Springer, 2001.

Kloeppel, James E. *Realm of the Long Eyes: A Brief History of Kitt Peak National Observatory*. San Diego: Univelt, 1983.

Knapp, Ralph E. *Geophysics*. Exeter, England: Pergamon Press, 1995.

Knauss, John. *Introduction to Physical Oceanography*. 2d ed. Long Grove, Ill.: Waveland Press, 2005.

Kolerstrom, Nicholas. *Newton's Forgotten Lunar Theory: His Contribution to the Quest for Longitude*. Santa Fe, N.Mex.: Green Lion Press, 2000.

Kosofsky, L. J., and Farouk El-Baz. *The Moon as Viewed by the Lunar Orbiter*. NASA SP-200. Washington, D.C.: Government Printing Office, 1970.

Kovalevsky, Jean. *Modern Astrometry*. New York: Springer, 1995.

Kowal, Charles T. *Asteroids: Their Nature and Utilization*. Chichester, England: Ellis Harwood, 1988.

Krisciunas, Kevin. *Astronomical Centers of the World*. Cambridge, England: Cambridge University Press, 1988.

Kroner, A., G. N. Hanson, and A. M. Goodwin, eds. *Archaean Geochemistry: The Origin and Evolution of the Archaean Continental Crust*. Berlin: Springer, 1984.

Krüger, Harald. *Jupiter's Dust Disc: An Astrophysical Laboratory*. Aachen, Germany: Shaker-Verlag, 2003.

Kuiper, Gerard P., and Barbara M. Middlehurst, eds. *Telescopes: Stars and Stellar Systems*. Chicago: University of Chicago Press, 1978.

Kump, Lee R., James Kasting, and Robert Crane. *The Earth System*. Upper Saddle River, N.J.: Prentice Hall, 2003.

Kundu, M. R., B. Woodgate, and E. J. Schmahl, eds. *Energetic Phenomena of the Sun*. Boston: Kluwer, 1989.

Kwok, Sun. *Physics and Chemistry of the Interstellar Medium*. New York: University Science Books, 2006.

Lambeck, Kurt. *The Earth's Variable Rotation: Geophysical Causes and Consequences*. New York: Cambridge University Press, 2005.

Lang, Kenneth R. *The Cambridge Guide to the Solar System*. Cambridge, England: Cambridge University Press, 2003.

_______. *Sun, Earth, and Sky*. 2d ed. New York: Springer, 2006.

Lang, Kenneth R., and Owen Gingerich, eds. *A Source Book in Astronomy and Astrophysics, 1900-1975*. Cambridge, Mass.: Harvard University Press, 1979.

Lankford, John, ed. *History of Astronomy: An Encyclopedia*. New York: Garland, 1997.

Lapedes, D. N., ed. *McGraw-Hill Encyclopedia of Geological Sciences*. New York: McGraw-Hill, 1978.

Lauretta, Dante S., and Harry Y. McSween, eds. *Meteorites and the Early Solar System II*. Tucson: University of Arizona Press, 2006.

Lemoine, M., J. Martin, and P. Peter, eds. *Inflationary Cosmology*. New York: Springer, 2008.

Leutwyler, Kristin, and John R. Casani. *The Moons of Jupiter*. New York: W. W. Norton, 2003.

Leverington, David. *Babylon to Voyager and Beyond: A History of Planetary Astronomy*. New York: Cambridge University Press, 2003.

Levin, Harold L. *The Earth Through Time*. 5th ed. Fort Worth: Saunders College Publishing, 1996.

Levine, Joel S., ed. *The Photochemistry of Atmospheres: Earth, the Other Planets, and Comets*. Orlando, Fla.: Academic Press, 1985.

Levinson, Alfred Abraham, ed. *Apollo 11 Lunar Science Conference: Proceedings*. 3 vols. Elmsford, N.Y.: Pergamon Press, 1970.

Levinton, Jeffrey S. *Genetics, Paleontology, and Macroevolution*. 2d ed. New York: Cambridge University Press, 2001.

Levy, David H. *Clyde Tombaugh, Discoverer of Planet Pluto*. New York: Sky, 2007.

_______. *David Levy's Guide to Observing Meteor Showers*. Cambridge, England: Cambridge University Press, 2008.

_______. *Impact Jupiter: The Crash of Comet Shoemaker-Levy 9*. New York: Basic Books, 2003.

_______. *The Quest for Comets: An Explosive Trail of Beauty and Danger*. New York: Plenum Press, 1994.

Lewis, Cherry. *The Dating Game: One Man's Search for the Age of the Earth*. Cambridge, England: Cambridge University Press, 2002.

Lewis, John S. *Physics and Chemistry of the Solar System*. 2d ed. San Diego, Calif.: Academic Press, 2004.

_______. *Rain of Iron and Ice: The Very Real Threat of Comet and Asteroid Bombardment*. New York: Basic Books, 1997.

Lewis, John S., and Ronald G. Prinn. *Planets and Their Atmospheres: Origin and Evolution*. New York: Academic Press, 1983.

Liddle, Andrew, and Jon Loveday. *The Oxford Companion to Cosmology*. New York: Oxford University Press, 2008.

Littmann, Mark. *Planets Beyond: Discovering the Outer Solar System*. New York: Dover, 2004.

Lockman, F. J., F. D. Ghigo, and D. S. Balsar, eds. *But It Was Fun: The First Forty Years of Radio Astronomy at Green Bank*. Washington, D.C.: National Radio Astronomy Observatory, 2007.

Lopes, Rosaly M. C., and T. K. P. Gregg. *Volcanic Worlds: Exploring the Solar System's Volcanoes*. New York: Springer, 2004.

Lopes, Rosaly M. C., and John R. Spencer. *Io After Galileo: A New View of Jupiter's Volcanic Moon*. Heidelberg: Springer, 2007.

Lorenz, Ralph, and Jacqueline Mitton. *Titan Unveiled: Saturn's Mysterious Moon Explored*. Princeton, N.J.: Princeton University Press, 2008.

Lovett, Laura, Joan Harvath, and Jeff Cuzzi. *Saturn: A New View*. New York: Harry N. Abrams, 2006.

Lowell, Percival H. *Mars and Its Canals*. New York: Macmillan, 1906.

Luisi, Pier Luigi. *The Emergence of Life: From Chemical Origins to Synthetic Biology*. New York: Cambridge University Press, 2006.

Lunar and Planetary Institute, Houston, Texas.

Basaltic Volcanism on the Terrestrial Planets. Elmsford, N.Y.: Pergamon Press, 1981.

Lyne, Andrew G., and Francis Graham-Smith. *Pulsar Astronomy*. 3d ed. Cambridge, England: Cambridge University Press, 2006.

McAnally, John W. *Jupiter, and How to Observe It*. New York: Springer, 2008.

McBride, Neil, and Gilmour Iain, eds. *An Introduction to the Solar System*. Cambridge, England: Cambridge University Press, 2004.

Maccarone, Thomas J. *From X-ray Binaries to Quasars: Black Holes on All Mass Scales*. New York: Kindle, 2006.

McConnell, Anita. *Geomagnetic Instruments Before 1900*. London: Harriet Wynter, 1980.

McElhinny, M. W., ed. *The Earth: Its Origin, Structure, and Evolution*. New York: Academic Press, 1979.

McFadden, Lucy-Ann Adams, Paul Robest Weissman, and T. V. Johnson, eds. *Encyclopedia of the Solar System*. San Diego: Academic Press, 2007.

Mackenzie, Dana. *The Big Splat: Or, How Our Moon Came to Be*. Hoboken, N.J.: John Wiley & Sons, 2003.

MacKenzie, Fred T. *Our Changing Planet: An Introduction to Earth System Science and Global Environmental Change*. Upper Saddle River, N.J.: Prentice Hall, 2002.

McLean, Ian S. *Electronic Imaging in Astronomy: Detectors and Instrumentation*. New York: Springer, 2008.

McSween, Harry Y., Jr. *Meteorites and Their Parent Planets*. 2d ed. New York: Cambridge University Press, 1999.

_______. *Stardust to Planets*. New York: St. Martin's Griffin, 1993.

Magli, Giulio. *Mysteries and Discoveries of Archaeoastronomy: From Giza to Easter Island*. New York: Springer, 2009.

Malphrus, Benjamin K. *The History of Radio Astronomy and the National Radio Astronomy Observatory: Evolution Toward Big Science*. Malabar, Fla.: Krieger, 1996.

Mammana, Dennis, and Donald McCarthy. *Other Suns, Other Worlds? The Search for Extrasolar Planetary Systems*. New York: St. Martin's Press, 1995.

Manuel, Oliver. *Origin of Elements in the Solar System: Implications of Post-1967 Observations.* New York: Springer, 2001.

Maran, Stephen P., ed. *The Astronomy and Astrophysics Encyclopedia*. Foreword by Carl Sagan. New York: Van Nostrand Reinhold, 1992.

Marov, Mikhail Ya, and David Grinspoon. *The Planet Venus*. New Haven, Conn.: Yale University Press, 1998.

Marschall, Laurence A. *The Supernova Story*. New York: Plenum Press, 1988.

Marten, Michael, and John Chesterman. *The Radiant Universe*. New York: Macmillan, 1980.

Mason, Brian. *Meteorites*. New York: John Wiley & Sons, 1962.

Mason, Brian, and William G. Melson. *The Lunar Rocks*. New York: Wiley-Interscience, 1970.

Meadows, A. J. *Early Solar Physics*. Elmsford, N.Y.: Pergamon Press, 1970.

Melchior, Paul. *The Earth Tides*. Oxford, England: Pergamon Press, 1966.

Melosh, H. J. *Impact Cratering: A Geologic Process*. New York: Oxford University Press, 1996.

Merrill, R. T., and M. W. McElhinney. *The Earth's Magnetic Field*. New York: Academic Press, 1983.

Meyer-Vernet, Nicole. *Basics of the Solar Wind*. Cambridge, England: Cambridge University Press, 2007.

Mezzacappa, Anthony, and George M. Fuller, eds. *Open Issues in Core Collapse Supernova Theory*. Hackensack, N.J.: World Scientific, 2005.

Michaud, Michael A. G. *Contact with Alien Civilizations: Our Hopes and Fears About Encountering Extraterrestrials*. New York: Springer, 2006.

Miller, Ron. *Extrasolar Planets*. Minneapolis, Minn.: Twenty-First Century Books, 2002.

_______. *Uranus and Neptune*. Brookfield, Conn.: Twenty-First Century Books, 2003.

Miller, S. L., and L. E. Orgel. *The Origins of Life on Earth*. Englewood Cliffs, N.J.: Prentice-Hall, 1974.

Milone, Eugene F., and William Wilson. *Solar System Astrophysics: Background Science on* *the Inner Solar System*. New York: Springer, 2008.

Miner, Ellis D. *Uranus: The Planet, Rings, and Satellites*. New York: Ellis Horwood, 1990.

Miner, Ellis D., and Randii R. Wessen. *Neptune: The Planet, Rings, and Satellites*. New York: Springer, 2002.

Misner, Charles W., Kip S. Thorne, and John A. Wheeler. *Gravitation*. San Francisco: W. H. Freeman, 1973.

Mitton, Simon. *Daytime Star: The Story of Our Sun*. New York: Charles Scribner's Sons, 1981.

Mobberley, Martin. *Supernovae and How to Observe Them*. New York: Springer, 2007.

Moche, Dinah L. *Astronomy: A Self-Teaching Guide*. 6th ed. New York: John Wiley & Sons, 2004.

Moldwin, Mark. *An Introduction to Space Weather*. Cambridge, England: Cambridge University Press, 2008.

Monod, Jacques. *Chance and Necessity*. Translated by Austryn Wainhouse. New York: Alfred A. Knopf, 1971.

Montesinos, Benjamín, Alvaro Giménez, and Edward F. Guinan, eds. *The Evolving Sun and Its Influence on Planetary Environments*. San Francisco: Astronomical Society of the Pacific, 2001.

Moore, Patrick. *Astronomical Telescopes and Observatories for Amateurs*. New York: W. W. Norton, 1973.

_______. *Guide to Mars*. New York: W. W. Norton, 1977.

_______. *Moore on Mercury: The Planet and the Missions*. New York: Springer, 2006.

_______. *On the Moon*. London: Cassell, 2001.

_______. *The Planet Neptune: An Historical Survey Before Voyager*. 2d ed. New York: John Wiley and Sons, 1996.

Morrison, David. *Voyages to Saturn*. NASA SP-451. Washington, D.C.: Government Printing Office, 1982.

Morrison, David, and Tobias Owen. *The Planetary System*. 3d ed. San Francisco: Pearson/Addison-Wesley, 2003.

Morrison, David, and Jane Samz. *Voyage to Jupiter*. NASA SP-439. Washington, D.C.: Government Printing Office, 1980.

Morrison, David, Sidney Wolf, and Andrew Fraknoi. *Abell's Exploration of the Universe*. 7th ed. Philadelphia: Saunders College Publishing, 1995.

Motz, Lloyd, ed. *Rediscovery of the Earth*. New York: Van Nostrand Reinhold, 1979.

Muller, Richard. *Nemesis: The Death Star*. New York: Weidenfeld & Nicolson, 1988.

Munk, W. H., and G. J. F. MacDonald. *The Rotation of Earth: A Geophysical Discussion*. New York: Cambridge University Press, 1960.

Murdin, Paul, and Lesley Murdin. *Supernova*. New York: Cambridge University Press, 1985.

Murray, Bruce, Michael C. Malin, and Ronald Greeley. *Earthlike Planets*. San Francisco: W. H. Freeman, 1981.

Mutch, Thomas A. *Geology of the Moon*. Rev. ed. Princeton, N.J.: Princeton University Press, 1972.

_______, comp. *The Geology of Mars*. Princeton, N.J.: Princeton University Press, 1976.

Nagy, B. *Carbonaceous Meteorites*. New York: Elsevier, 1975.

National Aeronautics and Space Administration. *The Case for Mars Concept Development for a Mars Research Station: Concept Development for a Mars Research Station*. San Francisco: University Press of the Pacific, 2002.

_______. *Preliminary Science Report: Apollo 11*. NASA SP-214. Washington, D.C: Government Printing Office, 1969.

_______. *Preliminary Science Report: Apollo 12*. NASA SP-235. Washington, D.C: Government Printing Office, 1970.

_______. *Preliminary Science Report: Apollo 14*. NASA SP-272. Washington, D.C.: Government Printing Office, 1971.

_______. *Preliminary Science Report: Apollo 15*. NASA SP-289. Washington, D.C.: Government Printing Office, 1972.

_______. *Preliminary Science Report: Apollo 16*. NASA SP-315. Washington, D.C.: Government Printing Office, 1972.

_______. *Preliminary Science Report: Apollo 17*. NASA SP-330. Washington, D.C.: Government Printing Office, 1973.

_______. *Viking 1, Early Results*. NASA SP-408. Springfield, Va.: National Technical Information Service, 1976.

National Aeronautics and Space Administration Advisory Council. *Earth System Science Overview*. Washington, D.C.: Government Printing Office, 1986.

National Research Council. *Changing Climate: Report of the Carbon Dioxide Assessment Committee*. Washington, D.C.: National Academy Press, 1983.

Newburn, R. L., M. Neugebauer, and Jurgen H. Rahe, eds. *Comets in the Post-Halley Era*. New York: Springer, 2007.

Nicolson, Iain. *The Sun*. New York: Rand McNally, 1982.

Nisbet, Evan G. *The Young Earth: An Introduction to Archaean Geology*. Winchester, Mass.: Unwin Hyman, 1987.

Noll, Keith S., and Harold A. Weaver, and Paul D. Feldman, eds. *The Collision of Comet Shoemaker-Levy 9 and Jupiter*. Cambridge, England: Cambridge University Press, 1996.

North, Gerald. *Observing Variable Stars, Novae, and Supernovae*. New York: Cambridge University Press, 2004.

North, John. *Cosmos: An Illustrated History of Astronomy and Cosmology*. Chicago: University of Chicago Press, 2008.

Norton, O. Richard. *The Cambridge Encyclopedia of Meteorites*. New York: Cambridge University Press, 2002.

_______. *Rocks from Space: Meteorites and Meteorite Hunters*. 2d ed. Missoula, Mont.: Mountain Press, 1998.

Norton, O. Richard, and Lawrence Chitwood. *Field Guide to Meteors and Meteorites*. London: Springer, 2008.

Novikov, Igor. *Black Holes and the Universe*. Translated by Vitaly Kisin. Cambridge, England: Cambridge University Press, 1995.

Noyes, Robert W. *The Sun, Our Star*. Cambridge, Mass.: Harvard University Press, 1982.

Ollivier, Marc, T. Encrenaz, F. Rocaves, and F. Selsis. *Planetary Systems: Detection, Formation, and Habitability of Extrasolar Planets*. New York: Springer, 2008.

Orloff, Richard W., and David M. Harland. *Apollo: The Definitive Sourcebook*. New York: Springer, 2006.

Oxlade, Chris. *Jupiter, Neptune, and Other Outer Planets*. New York: Rosen Central, 2007.

Ozima, Minoru. *The Earth: Its Birth and Growth*. Translated by J. F. Wakabayashi. New York: Cambridge University Press, 1981.

Pagel, Bernard. *Nucleosynthesis and Chemical Evolution of Galaxies*. Cambridge, England: Cambridge University Press, 1997.

Pannekoek, A. *A History of Astronomy*. London: Barnes & Noble Books, 1969.

Parker, Barry. *Einstein's Dream*. New York: Plenum Press, 1986.

Parker, Sybil P., ed. *McGraw-Hill Encyclopedia of Physics*. New York: McGraw-Hill, 1983.

Pasachoff, Jay M., and Will Tirion. *Field Guide to the Stars and Planets*. 5th ed. Boston: Houghton Mifflin, 1999.

Peek, Bertrand M. *The Planet Jupiter*. London: Macmillan, 1958.

Percy, John. *Understanding Variable Stars*. Cambridge, England: Cambridge University Press, 2007.

Peuker-Ehrenbrink, Bernhard, and Birger Schmitz, eds. *Accretion of Extraterrestrial Matter Throughout Earth's History*. New York: Kluwer Academic/Plenum, 2001.

Pickard, George L. *Descriptive Physical Oceanography: An Introduction*. 5th ed. New York: Pergamon Press, 1990.

Pomerantz, Martin A. *Cosmic Rays*. New York: Van Nostrand Reinhold, 1971.

Ponnamperuma, C., ed. *Cosmochemistry and the Origins of Life*. Dordrecht, Netherlands: Reidel, 1982.

Ponnamperuma, C., and A. G. W. Cameron. *Interstellar Communication: Scientific Perspectives*. Boston: Houghton Mifflin, 1974.

Prialnik, Dina. *An Introduction to the Theory of Stellar Structure and Evolution*. Cambridge, England: Cambridge University Press, 2000.

Pulsating Stars. 2 vols. Introductions by F. G. Smith, A. Hewish, and T. Gold. New York: Plenum Press, 1968-1969.

Rabinowitz, Avi. *Warped Spacetime, the Einstein Equations, and the Expanding Universe*. New York: Springer, 2009.

Raeburn, Paul, and Matt Golombek. *Un-*

covering the Secrets of the Red Planet. Washington, D.C.: National Geographic Society, 1998.
Raup, David M. *The Nemesis Affair*. New York: W. W. Norton, 1986.
Reid, Neil, and Suzanne Hawley. *New Light on Dark Stars: Red Dwarfs, Low-Mass Stars, Brown Stars*. 2d ed. New York: Springer Praxis, 2005.
Rey, H. A. *The Stars: A New Way to See Them*. 1952. Reprint. Boston: Houghton Mifflin, 1988.
Reynolds, Mike. *Falling Stars: A Guide to Meteors and Meteorites*. Mechanicsburg, Pa.: Stackpole Books, 2001.
Richardson, Robert S. *Exploring Mars*. New York: McGraw-Hill, 1954.
Ripley, S. Dillon. *Fire of Life: The Smithsonian Book of the Sun*. Washington, D.C.: Smithsonian Exhibition Books, 1981.
Robinson, Keith. *Spectroscopy: The Key to the Stars: Reading the Lines in Stellar Spectra*. New York: Springer, 2007.
Rosenburg, G. D., and S. K. Runcorn, eds. *Growth Rhythms and the History of the Earth's Rotation*. New York: John Wiley & Sons, 1975.
Rossi, Bruno. *Cosmic Rays*. New York: McGraw-Hill, 1964.
Rothery, David A. *Satellites of the Outer Planets: Worlds in Their Own Right*. New York: Oxford University Press, 1999.
Rucker, Rudolf. *Geometry, Relativity, and the Fourth Dimension*. New York: Dover, 1977.
Ruddiman, William F. *Earth's Climate: Past and Future*. 2d ed. New York: W. H. Freeman, 2008.
Russell, Chrisopher T. *The Cassini-Huygens Mission: Orbiter Remote Sensing Investigations*. New York: Springer, 2006.
_______. *Deep Impact Mission: Looking Beneath the Surface of a Cometary Nucleus*. New York: Springer, 2005.
Rybicki, George B., and Alan P. Lightman. *Radiative Processes in Astrophysics*. New York: Wiley, 1979.

Sagan, Carl. *Contact*. New York: Pocket, 1997.
_______. *Cosmos*. New York: Random House, 1980.
Sagan, Carl, and Ann Druyan. *Comet*. New York: Random House, 1985.
Salop, Lazarus J. *Geological Evolution of the Earth During the Precambrian*. Translated by V. P. Grudina. New York: Springer/Verlag, 1983.
Sartori, Leo. *Understanding Relativity: A Simplified Approach to Einstein's Theories*. Berkeley: University of California Press, 1996.
Savage, Candace. *Aurora: The Mysterious Northern Lights*. New York: Firefly Books, 2001.
Schaaf, Fred. *Comet of the Century: From Halley to Hale-Bopp*. New York: Copernicus, 1997.
Scharf, Caleb. *Extrasolar Planets and Astrobiology*. Herndon, Va.: University Science Books, 2008.
Schindewolf, Otto H. *Basic Questions in Paleontology: Geologic Time, Organic Evolution, and Biological Systematics*. Translated by Judith Schaefer. Chicago: University of Chicago Press, 1994.
Schlegel, Eric M. *The Restless Universe: Understanding X-ray Astronomy in the Age of Chandra and Newton*. New York: Oxford University Press, 2002.
Schmitt, Harrison J. *Return to the Moon: Exploration, Enterprise, and Energy in the Human Settlement of Space*. New York: Copernicus Books, 2006.
Schmude, Richard. *Uranus, Neptune, and Pluto and How to Observe Them*. New York: Springer, 2008.
Schneider, Peter. *Extragalactic Astronomy and Cosmology: An Introduction*. New York: Springer, 2006.
Schneider, Stephen E., and Thomas T. Arny. *Pathways to Astronomy*. 2d ed. New York: McGraw-Hill, 2008.
Schopf, J. William, ed. *Earth's Earliest Biosphere: Its Origin and Evolution*. Princeton, N.J.: Princeton University Press, 1984.
_______. *Life's Origin: The Beginnings of Biological Evolution*. Berkeley: University of California Press, 2002.
Schrunk, David, Burton Sharpe, Bonnie L. Cooper, and Madhu Thangavelu. *The Moon: Resources, Future Development, and Settlement*. New York: Springer Praxis, 2008.

Schultz, Peter H. *Moon Morphology: Interpretations Based on Lunar Orbiter Photography*. Austin: University of Texas Press, 1974.

Schwartz, Joseph, and Michael McGuinness. *Einstein for Beginners*. New York: Pantheon Books, 1979.

Schwinger, Julian. *Einstein's Legacy: The Unity of Space and Time*. New York: Scientific American Books, 1986.

Sears, D. W. *The Nature and Origin of Meteorites*. Bristol, England: Adam Hilger, 1978.

Seeds, Michael A. *Foundations of Astronomy*. 9th ed. Belmont, Calif.: Thomson Brooks/Cole, 2007.

_______. *Horizons: Exploring the Universe*. New York: Brooks/Cole, 2007.

Seibold, E., and W. Berger. *The Sea Floor*. New York: Springer, 1982.

Seielstad, George A. *At the Heart of the Web: The Inevitable Genesis of Intelligent Life*. Boston, Mass.: Harcourt Brace Jovanovich, 1989.

Sekanina, Zdenek, ed. *The Comet Halley Archive Summary Volume*. Pasadena, Calif.: Jet Propulsion Laboratory (International Halley Watch), California Institute of Technology, 1991.

Selley, Richard, Robin Cocks, and Ian Plimer, eds. *Encyclopedia of Geology*. 5 vols. Oxford, England: Elsevier Academic Press, 2005.

Serge, Brunier. *Solar System Voyage*. Translated by Storm Dunlop. New York: Cambridge University Press, 2000.

Serway, Raymond A., et al. *College Physics*. 7th ed. New York: Brooks/Cole, 2005.

Severny, A. *Solar Physics*. San Francisco: University Press of the Pacific, 2004.

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

Shklovskii, I. S., and Carl Sagan. *Intelligent Life in the Universe*. San Francisco: Holden-Day, 1966.

Short, Nicholas M. *Planetary Geology*. Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Shu, Frank H. *The Physical Universe: An Introduction to Astronomy*. Mill Valley, Calif.: University Science Books, 1982.

Silk, Joseph. *The Big Bang*. Rev. ed. New York: W. H. Freeman, 1989.

Sion, Edward M., Stephane Vennes, and Harry L. Shipman. *White Dwarfs: Cosmological and Galactic Probes*. New York: Springer, 2005.

Skinner, Brian J. *The Blue Planet: An Introduction to Earth System Science*. New York: John Wiley, 1995.

Skinner, Brian J., and S. C. Porter. *The Dynamic Earth*. 5th ed. New York: John Wiley & Sons, 2006.

Slade, Suzanne. *A Look at Jupiter*. New York: PowerKids Press, 2008.

Smart, William M. *The Origin of the Earth*. 2d ed. New York: Cambridge University Press, 1953.

Smith, David G., ed. *The Cambridge Encyclopedia of Earth Sciences*. New York: Cambridge University Press, 1982.

Smith, F. G. *Pulsars*. Cambridge, England: Cambridge University Press, 1977.

Smith, Michael D. *The Origin of Stars*. London: Imperial College Press, 2004.

Snow, Theodore P. *The Dynamic Universe*. Rev. ed. St. Paul, Minn.: West, 1991.

Sobel, Dava. *The Planets*. New York: Viking, 2005.

Sonett, C. P., M. S. Giampapa, and M. S. Matthews, eds. *The Sun in Time*. Tucson: University of Arizona Press, 1991.

Soon, Willie Wei-Hock, and Steven H. Yaskell. *The Maunder Minimum and the Variable Sun-Earth Connection*. Hackensack, N.J.: World Scientific, 2003.

Spangenburg, Ray, and Kit Moser. *A Look at Mercury*. New York: Franklin Watts, 2003.

_______. *A Look at Venus*. New York: Franklin Watts, 2002.

_______. *Meteors, Meteorites, and Meteoroids*. Secaucus, N.J.: Franklin Watts, 2002.

Spencer, John R., and Jacqueline Mitton, eds. *The Great Comet Crash: The Collision of Comet Shoemaker-Levy 9 and Jupiter*. Cambridge, England: Cambridge University Press, 1995.

Spitzer, Lyman, Jr. *Physical Processes in the Interstellar Medium*. New York: Wiley, 1998.

Spudis, Paul D. *The Geology of Multi-ring Impact Basins: The Moon and Other Planets*. New York: Cambridge University Press, 1993.

Squyres, Steve. *Roving Mars: Spirit, Opportu-*

nity, and the Exploration of the Red Planet. New York: Hyperion, 2006.
Stacey, F. D. *Physics of the Earth*. New York: John Wiley & Sons, 1977.
Stephenson, Bruce. *Kepler's Physical Astronomy*. Princeton, N.J.: Springer, 1994.
Stephenson, F. Richard. *Historical Eclipses and Earth's Rotation*. New York: Cambridge University Press, 2008.
Stern, Alan, and Jacqueline Mitton. *Pluto and Charon*. New York: Wiley, 1999.
Stix, Michael. *The Sun*. New York: Springer, 2004.
Stone, Robert G., Kurt W. Weiler, Melvyn L. Goldstein, and Jean-Louis Bouqueret, eds. *Radio Astronomy at Long Wavelengths*. New York: American Geophysical Union, 2000.
Strahler, Arthur N., and Alan H. Strahler. *Modern Physical Geography*. 4th ed. New York: John Wiley & Sons, 1992.
Strom, Robert G., and Ann L. Sprague. *Exploring Mercury: The Iron Planet*. New York: Springer, 2004.
Sullivan, W. T., ed. *The Early Years of Radio Astronomy*. New York: Cambridge University Press, 1984.
Sullivan, Walter. *Assault on the Unknown*. New York: McGraw-Hill, 1961.

Tabak, John. *A Look at Neptune*. London: Franklin Watts, 2003.
Taff, Laurence G. *Celestial Mechanics*. New York: Wiley-Interscience, 1985.
Tarbuck, Edward J., and Frederick K. Lutgens. *Earth: An Introduction to Physical Geology*. Illustrated by Dennis Tasa. 9th ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2008.
Tayler, Roger J. *Galaxies: Structure and Evolution*. Rev. ed. New York: Cambridge University Press, 1993.
Taylor, Edwin F., and John A. Wheeler. *Space-Time Physics*. San Francisco: W. H. Freeman, 1966.
Taylor, Stuart R. *Planetary Science: A Lunar Perspective*. Houston: Lunar and Planetary Institute, 1982.
Taylor, Stuart R., and Scott M. McLennan. *The Continental Crust: Its Composition and Evolution*. Boston: Blackwell Scientific, 1985.
Thackray, John. *The Age of the Earth*. New York: Cambridge University Press, 1989.
Thomas, Paul J., et al. *Comets and the Origin and Evolution of Life*. 2d ed. New York: Springer, 2006.
Thompson, Roy, and Frank Oldfield. *Environmental Magnetism*. London: Allen & Unwin, 1986.
Thornton, Stephen T., and Andrew Rex. *Modern Physics for Students and Engineers*. 3d ed. New York: Brooks/Cole, 2005.
Tielens, A. G. G. M. *The Physics and Chemistry of the Interstellar Medium*. Cambridge, England: Cambridge University Press, 2005.
Time-Life Books. *Comets, Asteroids, and Meteorites*. Alexandria, Va.: Author, 1990.
_______. *The Far Planets*. Alexandria, Va.: Author, 1988.
_______. *The Near Planets*. Alexandria, Va.: Author, 1989.
_______. *The New Astronomy*. Alexandria, Va.: Author, 1989.
Tipler, Paul A., and Ralph Llewellyn. *Modern Physics*. 5th ed. New York: W. H. Freeman, 2007.
Tobias, Russell R., and David G. Fisher, eds. *USA in Space*. 3d ed. Pasadena, Calif.: Salem Press, 2006.
Tocci, Salvadore. *A Look at Uranus*. New York: Franklin Watts, 2003.
Tombaugh, Clyde W., and Patrick Moore. *Out of the Darkness: The Planet Pluto*. New York: New American Library, 1981.
Trefil, James S. *The Moment of Creation: Big Bang Physics from Before the First Millisecond to the Present Universe*. New York: Charles Scribner's Sons, 1983.
_______. *Other Worlds: Images of the Cosmos from Earth and Space*. Washington, D.C.: National Geographic Society, 1999.
Trujillo, Alan, and Harold Thurman. *Essentials of Oceanography*. 9th ed. Upper Saddle River, N.J.: Prentice Hall, 2007.
Trumper, Joachin, and Gunther Hasinger. *The Universe in X Rays*. New York: Springer, 2008.
Tucker, Wallace. *The Star Splitters: The High Energy Astronomy Observatories*. NASA SP-466. Washington, D.C.: Government Printing Office, 1984.

Tucker, Wallace, and Riccardo Giacconi. *The X-ray Universe*. Cambridge, Mass.: Harvard University Press, 1985.

Tucker, Wallace, and Karen Tucker. *The Cosmic Inquirers: Modern Telescopes and Their Makers*. Cambridge, Mass.: Harvard University Press, 1986.

Tumlinson, Rick N., and Erin Medlicott, eds. *Return to the Moon*. New York: Collector's Guide, 2005.

Tyson, Neil deGrasse. *The Pluto Files*. New York: W. W. Norton, 2009.

Uchupi, E., and K. Emery. *Morphology of the Rocky Members of the Solar System*. New York: Springer, 1993.

Unsöld, Albrecht, and Bodo Baschek. *The New Cosmos: An Introduction to Astronomy and Astrophysics*. 5th ed. New York: Springer, 2001.

Van Allen, James A. *The Magnetospheres of Eight Planets and the Moon*. Oslo, Norway: Norwegian Academy of Science and Letters, 1990.

Van der Meer, Freek D., and Steven M. De Jong, eds. *Imaging Spectrometry*. New York: Kluwer Academic, 2002.

Van der Pluijm, Ben, and Stephen Marshak. *Earth's Structure*. 2d ed. New York: W. W. Norton, 2003.

Van Pelt, Michel. *Space Invaders: How Robotic Spacecraft Explore the Solar System*. New York: Springer, 2006.

Verschuur, Gerrit L. *Impact! The Threat of Comets and Asteroids*. New York: Oxford University Press, 1997.

_______. *The Invisible Universe: The Story of Radio Astronomy*. 2d ed. New York: Springer, 2006.

Vogel, Shawna. *Naked Earth: The New Geophysics*. New York: Plume, 1996.

Voit, Mark. *Hubble Space Telescope: New Views of the Universe*. New York: Harry N. Abrams, 2000.

Wagner, Jeffrey K. *Introduction to the Solar System*. Philadelphia: Saunders College Publishing, 1991.

Wall, J. V., ed. *Optics in Astronomy*. New York: Cambridge University Press, 1993.

Walter, Malcolm. *The Search for Life on Mars*. Cambridge, Mass.: Perseus Books, 1999.

Ward, Peter. *Life as We Do Not Know It: The NASA Search for (and Synthesis of) Alien Life*. New York: Penguin, 2007.

Ward, Peter, and Donald Brownlee. *Rare Earth: Why Complex Life Is Uncommon in the Universe*. New York: Springer, 2000.

Washburn, Mark. *Distant Encounters: The Exploration of Jupiter and Saturn*. New York: Harcourt Brace Jovanovich, 1982.

Wasson, John T. *Meteorites: Classification and Properties*. New York: Springer, 1974.

_______. *Meteorites: Their Record of Early Solar-System History*. New York: W. H. Freeman, 1985.

Webb, Stephen. *If the Universe Is Teeming with Aliens . . . Where Is Everybody? Fifty Solutions to Fermi's Paradox and the Problem of Extraterrestrial Life*. New York: Springer, 2002.

Wedepohl, Karl H. *Geochemistry*. New York: Holt, Rinehart and Winston, 1971.

Weeks, T. C. *Very High Energy Gamma Ray Astronomy*. New York: Taylor & Francis, 2003.

Weinberg, Steven. *Cosmology*. New York: Oxford University Press, 2008.

_______. *The First Three Minutes*. New York: Bantam Books, 1977.

Weiner, Jonathan. *Planet Earth*. New York: Bantam Books, 1986.

Weintraub, David A. *Is Pluto a Planet? A Historical Journey Through the Solar System*. Princeton, N.J.: Princeton University Press, 2006.

Wentzel, G. Donat. *The Restless Sun*. Washington, D.C.: Smithsonian Institution Press, 1989.

Wheeler, J. Craig. *Cosmic Catastrophes: Supernovae, Gamma-Ray Bursts, and Adventures in Hyperspace*. New York: Cambridge University Press, 2000.

Wheeler, John Archibald. *A Journey into Gravity and Spacetime*. New York: Scientific American Library, 1999.

Whipple, Fred L. *The Mystery of Comets*. Washington, D.C.: Smithsonian Institution Press, 1985.

White, Oran R., ed. *The Solar Output and Its Variation*. Boulder: Colorado Associated University Press, 1977.

Whitney, Charles. *The Discovery of Our Galaxy*. New York: Alfred A. Knopf, 1971.

Wicander, Reed, and James Monroe. *Historical Geology*. 5th ed. Florence, Ky.: Brooks/Cole, 2006.

Wilford, John Noble. *Mars Beckons: The Mysteries, the Challenges, the Expectations of Our Next Great Adventure in Space*. New York: Alfred Knopf, 1990.

Wilhelms, Don E. *The Geologic History of the Moon*. U.S. Geological Survey Professional Paper 1348. Washington, D.C.: Government Printing Office, 1987.

_______. *To a Rocky Moon: A Geologist's History of Lunar Exploration*. Phoenix: University of Arizona Press, 1994.

Wilkie, Tom, and Mark Rosselli. *Visions of Heaven: The Mystery of the Universe Revealed by the Hubble Space Telescope*. London: Hodder & Stoughton, 1998.

Will, Clifford. *Was Einstein Right? Putting General Relativity to the Test*. 2d ed. New York: Basic Books, 1993.

Woolfson, Michael. *The Formation of the Solar System: Theories Old and New*. London: Imperial College Press, 2007.

Wudka, Jose. *Space-Time, Relativity, and Cosmology*. New York: Cambridge University Press, 2006.

Wynn-Williams, Gareth. *The Fullness of Space: Nebulae, Stardust, and the Interstellar Medium*. Cambridge, England: Cambridge University Press, 1992.

Young, Carolynn, ed. *The Magellan Venus Explorers' Guide*. Pasadena, Calif.: Jet Propulsion Laboratory, California Institute of Technology, National Aeronautics and Space Administration, 1990.

Young, Hugh D., and Roger A. Freedman. *University Physics with Modern Physics*. 11th ed. New York: Addison Wesley, 2003.

Zeilik, Michael. *Astronomy: The Evolving Universe*. 9th ed. New York: John Wiley and Sons, 2002.

Zeilik, Michael, and Stephen A. Gregory. *Introductory Astronomy and Astrophysics*. 4th ed. Fort Worth, Tex.: Saunders College Publishing, 1998.

Zimmerman, Robert. *The Universe in a Mirror: The Saga and the Visionaries Who Built It*. Princeton, N.J.: Princeton University Press, 2008.

Zirker, J. B. *An Acre of Glass: A History and Forecast of the Telescope*. Baltimore: Johns Hopkins University Press, 2005.

_______. *Journey from the Center of the Sun*. Princeton, N.J.: Princeton University Press, 2002.

_______. *Sunquakes: Probing the Interior of the Sun*. Baltimore: Johns Hopkins University Press, 2003.

_______. *Total Eclipses of the Sun*. Expanded ed. Princeton, N.J.: Princeton University Press, 1995.

Zubrin, Robert. *Entering Space: Creating a Spacefaring Civilization*. New York: Tarcher, 2000.

Web Sites

Listed below are more than 110 authoritative Web sites current as of 2009. Although every effort has been made to ensure accuracy, Web sites are continually being updated, and there may be changes to the site address listed. A subject or keyword search through any of the major search engines will help locate a new address if the universal resource locators (URLs) listed below have been changed.

Adler Planetarium
http://www.adlerplanetarium.org

Amazing Space
http://amazing-space.stsci.edu

American Meteor Society
http://www.amsmeteors.org

Ames Research Center
http://www.nasa.gov/centers/ames

Apollo Program
http://spaceflight.nasa.gov/history/apollo

Astronomical Society of the Pacific
http://www.astrosociety.org/index.html

Astronomy Café, The
http://www.astronomycafe.net

Astronomy Magazine
http://www.astronomy.com

Astronomy Now Magazine
http://www.astronomynow.com

Astronomy Picture of the Day
http://antwrp.gsfc.nasa.gov/apod

Beijing Planetarium
http://www.bjp.org.cn/en/index.htm

British National Space Centre
http://www.bnsc.gov.uk

Cambridge Cosmology Public Home Page
http://www.damtp.cam.ac.uk/user/gr/public/cos_home.html

Canadian Space Agency/L'Agence Spatiale Canadienne
http://www.space.gc.ca/asc/index.html

Cassini Equinox Mission
http://saturn.jpl.nasa.gov

Cassini-Huygens Mission
http://www.nasa.gov/mission_pages/cassini/main

Chandra X-ray Observatory Center
http://chandra.harvard.edu

Comet Shoemaker-Levy Collision with Jupiter
http://www2.jpl.nasa.gov/sl9

Compton Gamma Ray Observatory
http://cossc.gsfc.nasa.gov

Dawn Mission
http://dawn.jpl.nasa.gov

Deep Impact
http://deepimpact.jpl.nasa.gov

Discovery Program
http://discovery.nasa.gov

Earth Observations Photography, Space Shuttle
http://earth.jsc.nasa.gov/sseop/efs

Exoplanets: The Search for Extrasolar Planets
http://exoplanets.org

Exploration System Mission Directorate
http://exploration.nasa.gov

Explorer Missions
http://nssdc.gsfc.nasa.gov/multi/explorer.html

Galileo: Journey to Jupiter
http://www2.jpl.nasa.gov/galileo

Gamma-ray Large Area Space Telescope
http://glast.gsfc.nasa.gov

Gemini Program
http://www-pao.ksc.nasa.gov/kscpao/history/gemini/gemini.htm

Goddard Space Flight Center
http://www.nasa.gov/centers/goddard

Gravity Probe B
http://www.gravityprobeb.com

Great Images in NASA
http://grin.hq.nasa.gov

Griffith Park Observatory
http://www.griffithobs.org

Harvard-Smithsonian Center for Astrophysics
http://www.cfa.harvard.edu

Hayden Planetarium
http://www.haydenplanetarium.org

Heavens-Above Satellite Observations
http://www.heavens-above.com

Hubblesite News Center
http://hubblesite.org/newscenter

International Year of Astronomy: 2009
http://www.astronomy2009.org

Jet Propulsion Laboratory (JPL)
http://www.jpl.nasa.gov

Jodrell Bank Centre for Astrophysics
http://www.jb.man.ac.uk

Johannesburg Planetarium
http://www.planetarium.co.za/

Johnson Space Center
http://www.nasa.gov/centers/johnson/home/index.html

JPL: Goldstone Complex
http://deepspace.jpl.nasa.gov/dsn/gallery/goldstone.html

JPL: Multimedia
http://www.jpl.nasa.gov/multimedia/index.cfm

JPL: The Solar System
http://www.jpl.nasa.gov/solar-system/index.cfm

JPL: Solar System Dynamics
http://ssd.jpl.nasa.gov

JPL: Stars and Galaxies
http://www.jpl.nasa.gov/stars-galaxies/index.cfm

Kennedy Space Center
http://www.nasa.gov/centers/kennedy

Lunar and Planetary Institute
http://www.lpi.usra.edu

Lunar Prospector
http://lunar.arc.nasa.gov

Mars Exploration Program
http://mars.jpl.nasa.gov

Mars Exploration Rover Mission
http://marsrovers.jpl.nasa.gov/home

Mars Global Surveyor: Mars Orbiter Camera
http://mars.jpl.nasa.gov/mgs/msss/camera/images/

Mars Today
http://www-mgcm.arc.nasa.gov

Marshall Space Flight Center
http://www.nasa.gov/centers/marshall

Melbourne Planetarium
http://museumvictoria.com.au/planetarium

National Aeronautics and Space Administration (NASA)
http://www.nasa.gov

NASA: Images
http://www.nasaimages.org

NASA: Multimedia
http://www.nasa.gov/multimedia/index.html

NASA: Origins of the Universe
http://origins.jpl.nasa.gov

NASA: Planetary Photojournal
http://photojournal.jpl.nasa.gov

NASA: Science
http://www.earth.nasa.gov

NASA: Solar System Mission
http://www.nasa.gov/missions/timeline/current/solar-system_missions.html

NASA Space Place
http://spaceplace.nasa.gov/en/kids

NASA Space Place: "Ask Dr. Marc"
http://spaceplace.nasa.gov/en/kids/phonedrmarc

NASA: Space Science
http://spacescience.nasa.gov

NASA: Space Science Data Center
http://nssdc.gsfc.nasa.gov

NASA: Space Science Photo Gallery
http://nssdc.gsfc.nasa.gov/photo_gallery

NASA: Space Science Planetary Missions
http://nssdc.gsfc.nasa.gov/planetary/projects.html

NASA: Sun-Earth Connection Information Forum
http://sunearth.gsfc.nasa.gov/missions/index.php

National Astronomy and Ionosphere Center: Arecibo Observatory
http://www.naic.edu

National Geographic Astronomy
http://www.nationalgeographic.com/stars

National Geographic's Virtual Solar System
http://science.nationalgeographic.com/science/space/solar-system

National Oceanic and Atmospheric Administration
http://www.noaa.gov

National Optical Astronomy Observatory
http://www.noao.edu

National Radio Astronomy Observatory
http://www.nrao.edu

National Weather Service
http://www.nws.noaa.gov

New Horizons Pluto-Kuiper Belt Mission
http://pluto.jhuapl.edu

The Nine Planets: A Multimedia Tour of the Solar System
http://seds.lpl.arizona.edu/nineplanets/nineplanets

Pacific Science Center
http://www.pacsci.org

Phoenix Mars Lander 2007
http://phoenix.lpl.arizona.edu

Planetary Society, The
http://planetary.org

Planet Quest: Exoplanet Exploration
http://planetquest.jpl.nasa.gov

PlanetScapes
http://planetscapes.com

Russian Space Web: Chronology of Space Exploration
http://www.russianspaceweb.com/chronology.html

Satellite Information, World Data Center System
http://www.ngdc.noaa.gov/wdc

Search for Extraterrestrial Intelligence Institute
http://www.seti-inst.edu

Sky and Telescope Magazine
http://www.skyandtelescope.com

Solar and Heliospheric Observatory
http://sohowww.nascom.nasa.gov

Solar System Exploration
http://sse.jpl.nasa.gov

Space Foundation
http://www.spacefoundation.org

Space Station: Science Operations News
http://www.scipoc.msfc.nasa.gov

Space Telescope Science Institute
http://www.stsci.edu

Space Telescope Science Institute: Hubble Space Telescope
http://www.stsci.edu/hst

Space Telescope Science Institute: James Webb Space Telescope
http://www.stsci.edu/jwst

Space.com: Astronomy and Science News and Information
http://www.space.com/scienceastronomy

Spaceref.com
http://www.spaceref.com

Spaceweather.com
http://www.spaceweather.com

Spitzer Space Telescope
http://www.spitzer.caltech.edu/spitzer

Stanford Solar Center
http://solar-center.stanford.edu

Stardust Project
http://stardust.jpl.nasa.gov

Swift Gamma Ray Burst Mission
http://swift.gsfc.nasa.gov/docs/swift/swiftsc.html

The Two Micron All Sky Survey
http://pegasus.phast.umass.edu

Ulysses
http://ulysses.jpl.nasa.gov

Ulysses: European Space Agency
http://sci.esa.int/ulysses

U.S. Geological Survey
http://www.usgs.gov

Views of the Solar System
http://www.solarviews.com/eng/homepage.htm

W. M. Keck Observatory
http://www.keckobservatory.org

Welcome to the Planets
http://pds.jpl.nasa.gov/planets

Wilkinson Microwave Anisotropy Probe
http://map.gsfc.nasa.gov

Subject Index

Abbot, Charles Greeley, 752, 784
Abell, George, 884
Absolute magnitude, 957
Absolute space, 957
Absolute temperature scale, 957
Absolute time, 957
Absolute zero, 957
Absorption lines, 957
Absorption spectra, 201, 299, 303, 552, 560, 666, 714, 717, 730, 736, 748, 751, 957
Acasta gneiss, 106, 117
Acceleration, 957; of gravity, 957. *See also* Gravitational acceleration
Accretion, 104, 108, 576, 579, 618, 769, 774, 776, 906, 957; homogeneous, 140; mountain building, 132; water formation, 168
Accretion disk, 957
Achilles (asteroid), 8
Achondrites, 143, 468-472, 957; SNCs, 492
Acquisition, 957
Active cavity radiometer irradiance monitor, 754
Active Magnetospheric Particle Tracer Explorers, 790
Active optics, 826
Adams, Fred, 666
Adams, John Couch, 497, 527, 531
Adams, Walter Sydney, 944
Adams ring (Neptune), 532
Adaptive optics, 826
Adrastea (moon), 341, 608
Aerodynamics, 957
Aerography, 957
Aeronautics, 957
Aeronomy, 957
Aerospace, 957
Agonic line, 150
Ahnighito meteorite, 484
Aine (Venus), 939
Airglow, 957
Airy, Sir George, 527
Alais (meteorite), 473
Alba Patera (Mars), 443
Albedo, 260, 957
Alcock, E., 292
Aldebaran (star), 270, 672
Aldrin, Edwin E. “Buzz,” Jr., 384, 391
Alfvén, Hannes, 581
ALH 84001 (meteorite), 235, 493
Allen Telescope Array, 712
Allende (meteorite), 474-475
Almost stars. *See* Brown dwarf stars
Alpha Carinae (star), 822
Alpha decay, 106
Alpha Orionis (star), 672, 822
Alpha particles, 957
Alpha Regio (Venus), 930
Alpha Tauri (star), 672
Alpher, Ralph A., 25
Alphonsus (lunar crater), 362, 366
ALSEPs. *See* Apollo Lunar Science Experiment Packages
Alt-azimuth coordinate system, 66
Altair Lunar Lander, 372
Altitude, 957
Aluminum, 768
Alvarez, Luis, 11, 283, 513, 514
Alvarez, Walter, 11, 284, 513-514
Amalthea (moon), 341, 345; Jupiter’s magnetosphere, 338
American Geophysical Union, 129
Amino acids, 110, 232, 356, 856; meteorites, 473-474
Ammonia, 774
Amor asteroids, 8, 13; Eros, 10
Amphibole, 116
AMPTEs. *See* Active Magnetospheric Particle Tracer Explorers
Anaerobic life, 233
Anaxagoras, 495
Anderson, John D., 336
Andesite, 957
Anemometer, 958
Angstrom, 958
Angular momentum, 597, 655, 958; solar-system distribution, 777; solar-system formation, 774
Angular velocity, 598
Annihilation, 869
Annular solar eclipse, 193
Anorthosites, 369, 958; lunar, 369, 371
Antarctica, 119; meteorite finds, 491
Anthropic principle, 83, 872, 958
Antimatter, 80, 199, 869, 953, 958
Antiparticles, 958
Antoniadi, Eugene M., 435
Aphelion, 96, 178, 596, 958
Aphrodite Terra (Venus), 930-931
Apian, Peter, 48
Apoapsis, 958
Apocynthion, 958
Apogee, 958
Apollo-Amor asteroids, 281
Apollo asteroids, 8-10, 14, 16
Apollo Lunar Science Experiment Packages, 375
Apollo missions, 364, 619; experiments, 395; Hadley Rille, 379; heat-flow experiments, 374; rock and soil samples, 92, 119, 234,

376, 384, 391, 491, 628; seismological studies, 587
Apollo Telescope Mount, 719, 779, 863
Apollonius of Perga, 599
Apolune, 958
Apophis (asteroid), 12
Apparent magnitude, 958
Apparent motion, 958
Apparent size, 958
Aquarid meteor showers, 497
Arago ring (Neptune), 532
Archaebacteria, 233
Archaeoastronomy, 1-6
Archean eon, 117, 119, 958; rocks, 119
Arecibo radio telescope, 628, 828
Argon, 769
Argus project, 912
Argyre (Mars), 422
Ariel (moon), 704, 902
Aristarchus, 599
Aristotle, 495, 599
Armalcolite, 391
Armstrong, Neil, 391
Arp 220 galaxy, 291
Arsia Mons (Mars), 423, 448
Artificial satellite, 958
Asaro, Frank, 514
Ascension, 66
Ascraeus Mons (Mars), 448
Asteroid belt, 7, 576, 599, 602, 621, 959; formation, 774; orbital resonance, 603
Asteroids, 6-17, 958; Amor group, 8, 13-14; Apollo group, 8-10, 14, 16; C-type, 9; carbonaceous, 439, 961; Ceres, 40-44; Charon as, 646; Chiron, 8; D-type, 9; E-type, 9; Earth impacts, 599; economic potential, 16; Eros, 41; Hidalgo, 8; Hilda, 7; interiors, 584; Itokawa, 41; Juno, 7, 9; M-type, 9; mass extinctions, 513; origins, 15, 438, 776; Pallas, 7; and planet formation, 43; S-type, 9; threats to Earth, 12, 16, 283; Trojan group, 9-10, 603; U-type, 9; Vesta, 7, 41
Asthenospheres, 130, 135, 163, 165, 186, 188, 959; lunar, 375; Mercury, 461
Astro 1, 864
Astrobiology, 216. *See also* Exobiology
Astroblemes, 167, 279-286, 959; Canada, 280; Sahara Desert, 283
Astrometry, 228
Astron satellite, 864
Astronautics, 959
Astronomical coordinate systems, 65-69, 180
Astronomical unit (AU), 959
Astronomy, 959
Astrophysics, 959. *See also* Big bang; Cosmology; Gamma-ray astronomy; Infrared astronomy; Ultraviolet astronomy; X-ray astronomy
Atacama Large Millimeter/Submillimeter Array, 829
Ataxites, 481, 959
Aten (Io), 307
Aten asteroids, 9, 12
Atlas (moon), 707
Atmospheres, 564-571, 959; captured, 564; early, 770; Earth, 95, 108-114, 117, 167-168, 178, 565, 679, 721; evolution of life, 110; Ganymede, 245; Jovian planets, 311, 770; Jupiter, 318-330, 568, 632, 679; life's effect on, 232; Mars, 415-421, 566, 632, 639; Mercury, 460; Moon, 619; Neptune, 313, 516-520, 568; outgassing, 108, 564; Pluto, 643; Saturn, 568, 674-682, 689; solar, 713-723, 764; terrestrial planets, 415; Titan, 852, 854-855; Triton, 858-859; Uranus, 313, 568, 886-890; Venus, 565, 631, 637, 639, 914-922, 932, 934
Atmospheric Cherenkov Telescopes, 951
Atmospheric pressure, 959
Atom, 959
Atomic clocks, 181
Atomic number, 959
Attitude, 959
AU. *See* Astronomical unit
AU Microscopium (star), 667
Aubrey, John, 1
Aubrites, 470
Auroral kilometric radiation, 21
Auroral oval, 160
Auroras, 590, 959; australis, 959; borealis, 959; Earth, 17-23, 72, 156, 161, 297, 591, 728-729, 788, 790, 910, 912-913, 916; Jupiter, 21; Mars, 21; Neptune, 21, 529; Saturn, 21, 689-690; Uranus, 21, 894
Australia, 119
Australia Telescope, 829
Autumnal equinox, 67
Aveni, Anthony, 5
Avionics, 959
AX Microscopium, 666
Axis, 959
Axis tilt, 959
Azimuth, 960

Baade, Walter, 662
Babcock, Harold, 746
Babcock, Horace W., 743-744, 746, 813
Bahcall, John N., 540
Baily's beads, 195
Baity, Elizabeth Chesley, 1
Baldwin, Ralph, 366
Ballistics, 960
Balloon Observations of Millimetric Extragalactic Radiation and Geomagnetics, 26, 82, 881
Barnard's star, 67, 220, 227, 666, 725

Barringer, Daniel, 497
Barringer crater, 280
Bartels, J., 788
Barycenters, 90, 597-598, 960
Baryons, 960
Basalts, 186, 370, 469, 770, 960; lunar, 369, 371, 378, 391; Venus, 926-928, 933
Basins, 960. *See also* Craters; Impact craters; Maria
BATSE. *See* Burst and Transient Source Experiment
Beebe, Reta, 321
Bell, Jocelyn, 653, 662
Bellerophon. *See* 51 Pegasi
Benzenberg, J. F., 495
BeppoSax, 951
Berthollet, Claude Louis, 577
Bessel, Friedrich Wilhelm, 882
Beta decay, 76, 106, 538, 960
Beta Orionis (star), 822
Beta Regio (Venus), 930-931
Beta Taurid meteor showers, 499
Betelgeuse (star), 270, 672, 822
Bethe, Hans Albrecht, 25, 726, 851
Biermann, Ludwig, 788
Big bang, 24-31, 79-80, 544, 546, 808, 821, 867, 875; element formation, 767; relativity theory, 251
Big bang theory, 875, 960
Big crunch, 30, 800, 960
Bighorn Medicine Wheel, 4
Binary stars, 193, 220, 303, 544, 546, 548, 960; Sun's companion, 556; supernovae, 818; ultraviolet research, 865
Biosatellites, 960
Biot, Jean-Baptiste, 495
Biotelemetry, 960
Birefringence, 561
Black dwarf stars, 272, 551, 671, 724, 807, 849, 942-946, 960
Black holes, 544, 547, 798-799, 960; as burster sources, 241; galactic centers, 28; Milky Way, 501-502, 504, 546; mini, 251; neutron stars and, 546, 850; relativity and, 243; stellar-mass, 250; super-massive, 251; supernova remnants, 272, 544-545, 807; term coinage, 819; theory of, 250; as X-ray emitters, 546
Black smokers, 357
Blackbodies, 737, 752
Blackbody radiation, 197, 201, 552, 871
Blackett, Patrick M. S., 151
Blue giant stars, 542, 544
Blue-green algae. *See* Cyanobacteria
Blue supergiant stars; supernovae, 820. *See also* Supergiant stars
Blueberries, 455
Blueshifting, 561, 960. *See also* Doppler shifts
Bode, Johann, 6, 40, 572
Bode's law, 40
Body waves, 164
Bogard, Donald D., 790
Bohr, Niels, 726
Bolometer, 293, 752, 784-785
Bolometric magnitude, 960
Boltzmann, Ludwig, 201
Bondi, Hermann, 29, 875
BOOMERANG. *See* Balloon Observations of Millimetric Extragalactic Radiation and Geomagnetics
Borrelly's comet, 51, 60
Boss, Alan, 668
Bow shock effect, 159, 590, 960
Boynton, William, 457
Brachina, 471
Bradley, James, 616
Brahe, Tycho, 595, 616, 817
Braille (asteroid), 60
Brandes, H. W., 495
Branes, 83
Brans, Carl, 732, 760
Breccias, 281, 961; impact, 971; lunar, 371, 391
Bremsstrahlung radiation, 18, 23, 198, 947
Brown, Mike, 87, 210
Brown dwarf stars, 31-34, 59, 62, 220, 230, 291, 544, 961
Brownlee, Donald, 61
Brunini, Adrian, 907
Bruno, Giordano, 227
Buffon, Comte de, 580
Bullard, Sir Edward Crisp, 151
Burbidge, Geoffrey, 77, 553, 809
Burbidge, Margaret, 77, 553, 809
Burke, Bernard, 316, 335
Burst and Transient Source Experiment, 239, 951
Bursters, 238-244, 543-549
Butler, Paul, 222-223, 225

C-type asteroids, 9
Cabot, Sebastian, 150
Calcium, 768
Calderas, 961; Io, 305; lunar, 362; Mars, 448; Venus, 939
Calendars, 68
Callisto (moon), 35-40, 213, 218, 309, 314, 345-351, 621, 701; craters, 214; density, 218, 579; surface, 347; water, 349
Caloris Basin (Mercury), 461-462, 602, 630, 961
Calypso (moon), 603
Cameron, Alistair, 553
Canals (Mars), 401
Candles, 876, 945
Canopus (star), 822, 961
Canyon Diablo, 497
Carbon-based life, 266
Carbon dioxide, 961; atmospheric, 109, 261; global warming, 262; Venus, 915
Carbon 14, 745

Carbon monoxide on Neptune, 518
Carbon-nitrogen-oxygen cycle, 726, 849
Carbonaceous chondrites, 167, 355-356, 358, 472-478, 578
Carbonates, 109
Carrington, Richard C., 17, 729, 792, 811
Cassegrain telescopes, 961
Cassen, Patrick, 587
Cassini, Gian Domenico, 55, 275, 325, 428, 603, 605, 607, 624, 698, 708
Cassini division, 603-605, 607, 608, 694, 698, 961
Cassini mission, 235, 315, 321, 681; auroras, 21; Enceladus, 205, 314, 629; Great Red Spot, 327; Iapetus, 276; Jupiter, 341; Masursky asteroid, 14; Saturn, 678, 684; solar radio bursts, 757; Titan, 216, 853
Catastrophic hypothesis, 580
Catastrophism, 961
Catenae (Ganymede), 246
Cavendish, Henry, 253
Cavendish balance, 253
Cehtzuc, 3
Celestial equator, 66, 961
Celestial mechanics, 961
Celestial poles, 961
Celestial sphere, 961
Celsius scale, 961
Centaurs, 85, 353
Centrifugal force, 961
Centrifuge, 961
Cepheid variable stars, 874
Cepheid variables, 544, 838, 961; as standard candles, 877
Ceres (dwarf planet), 6-7, 9, 11, 40-44, 85, 572; Dawn mission, 15
CFCs. *See* Chlorofluorocarbons
CGRO. *See* Compton Gamma Ray Observatory
Chamberlin, Thomas Chronder, 581
Chandler's wobble, 180
Chandra X-Ray Observatory, 545, 547, 834, 837, 840, 951; auroras, 21
Chandrasekhar, Subrahmanyan, 547, 809, 944
Chandrasekhar limit, 818, 944, 962
Channeled Scabland (Washington State), 443
Chao Meng-Fu (Mercury), 464
Chapman, Sydney, 788
Charge, 962
Charge-coupled devices, 559, 949
Charged particles, 589, 743
Charon (dwarf planet), 85, 352, 597, 615, 621-622, 624, 634, 641-649; interior, 585; mass, 598
Chassignites, 470
Chauncy, Henry, 1
Chemical elements, 767-773, 966; creation of, 24-25, 75, 808, 818, 821; rare, 772. *See also* Nucleosynthesis
Chemical evolution, 962
Cherenkov counters, 75
Cherenkov radiation, 199, 538, 962
Cherenkov telescopes, 199
Cherts, 355, 962
Chichén Itzá, 3
Chicxulub Crater, 284, 512, 515
Chiron (asteroid), 8, 58, 85, 353
Chladni, Ernst, 495
Chlorofluorocarbons, 261, 263
Chondrites, 9, 143, 295, 468, 477-480, 583, 769, 776, 962; carbonaceous, 355-356, 472-478, 961; composition, 771; enstatite, 967; ordinary, 979
Chondrules, 468, 473, 477, 578, 962; porphyritic, 981; radiating pyroxene, 982
Christy, James W., 352, 624, 642, 644, 647
Chromosphere, 69, 713-717, 736-737, 757, 764, 780, 792, 962; solar flares, 727-732; ultraviolet observations, 780
Churyumov-Gerasimenko (comet), 51
Circle of illumination, 962
Circumstellar disk detection method, 228
CKBOs. *See* Kuiper Belt objects
Clairaut, Alexis-Claude, 46
Clark, Alvan Graham, 942
Clay minerals, 358
Clementine mission, 364, 372, 376, 382, 387, 619
Climate, 962
Climate change, 260-265, 420, 745; Earth, 98; models, 263; solar impact, 750, 752, 784-787, 814
Clocks, 181
Closed universe, 78, 875, 962
Cloud feedback, 263
Clouds; Jupiter, 568; Mars, 567; Neptune, 517, 569; Saturn, 568, 676, 679, 683; Titan, 852; Triton, 858; Uranus, 568, 888-889; Venus, 566, 631, 915, 922, 929, 938
Clusters and superclusters, 26, 29, 78, 81, 292, 301, 504, 874, 879-886, 968, 988. *See also* Galaxies
CMB. *See* Core-mantle boundary
CMBR. *See* Cosmic background radiation
CMEs. *See* Coronal mass ejections
CNO cycle. *See* Carbon-nitrogen-oxygen cycle
Coal, 118

Coalsack (nebula), 503
COBE. *See* Cosmic Background Explorer
Cochran, William, 224
Cocke, W. John, 653
Cocoon nebula, 575
Coesite, 962
Colbert, Jean-Bapiste, 607
Collimation methods, 947
Collins, Michael, 391
Collins, Robin Ashley, 653
Colombo, Giuseppe, 466
Colonization of the Moon, 382
Color-magnitude diagram, 269
Colorimetry, 13
Colors, 752
Columbus, Christopher, 150
Comas (cometary), 58, 555, 962
Comet period, 962
Comet Rendezvous and Asteroid Flyby, 51, 60
Comet showers, 62
Comets, 56-65, 572, 790, 962; Borrelly, 51, 60; Chiron, 58; coma, 58; composition, 771; Donati's of 1858, 48; Encke, 499; formation, 774, 776; Great Comet of 1843, 48; Great Daylight of 1910, 49; Great September of 1882, 48; Grigg-Skjellerup, 51; Halley's, 44-52, 56-58, 62, 233, 497; historical sightings, 46; interiors, 45; IRAS-Araki-Alcock, 292; Kohoutek, 57; Kuiper Belt, 58-59, 973; long-period, 555; lunar impacts, 366; motion of, 44; nuclei, 44, 46, 58; orbits, 46; origins, 58-59, 352-354, 555-558, 579; rotation, 45; Shoemaker-Levy 9, 52-56; short-period, 555; surface features, 46; tails, 58; ultraviolet images, 865; water, 46, 167; Wild 2, 233. *See also specific names*
Comparative planetology, 570, 630-637; terrestrial, 637-641
Compasses, 149
COMPTEL. *See* Imaging Compton Telescope
Compton, Arthur, 239
Compton Gamma Ray Observatory, 239, 242, 834, 836, 840, 951
Compton scattering, 239, 739, 949
Condensation (planet formation), 576, 962
Condensation temperature, 962
Conjunction, 962
Conservation law, 962
Conservation of angular momentum, 962
Constellation, 962
Constellation program, 372-373, 382, 393, 435, 623
Constellations, 180
Continental crust, 132, 136, 142, 186, 770
Continental drift, 150, 153, 587, 625, 628, 752
Convection, 963; Earth, 122, 145, 583; solar, 715, 733, 740, 743-744, 749, 756, 764, 813; stellar, 398
Convective equilibrium, 963
Convergent plate boundaries, 130
Coorbital resonance, 602
Coorbital satellites, 963
Coordinate systems, 65-69, 180
Copernicus (satellite), 863
Copernicus, Nicolaus, 92, 177, 227, 572, 599, 616
Cordelia (moon), 899
Core-collapse supernovae, 818, 849
Core-diffracted phases, 963
Core-mantle boundary, 121, 123, 125-129, 187-188
Core-mantle transition zone, 125
Core-reflected phases, 963
Core samples, 963; lunar, 386-387
Core-transmitted phases, 963
Cores, 963; stellar, 987
Cores of planets; Earth, 115, 121-125, 145, 151, 159, 374, 582; Jovian, 311-312, 635; Mercury, 460; meteoritic, 471; Moon, 374-378, 391; Neptune, 525, 529; Saturn, 675, 683-686, 691; Uranus, 890-891, 894; Venus, 932
Cores of stars, 544, 550-552; red giants, 671; red supergiants, 671; Sun, 739-742, 746, 764
Coriolis, Gaspard-Gustave de, 616
Coriolis effect, 146, 151, 178, 181, 337, 613, 616, 963
Corner reflector arrays, 396
Corona (Sun's), 589, 714, 717-723, 737, 765, 779-780, 963; holes, 69-73; mass ejections, 69-73
Coronae (Venus), 931, 939
Coronagraph, 718, 963
Coronal holes, 69-73, 719-720, 789, 792
Coronal loops, 70
Coronal mass ejections, 69-73, 297, 721, 745, 758; dangers from, 298
Coronographs, 715
COROT mission, 841
Corrective Optics Space Telescope Axial Replacement, 289
Cosmic Background Explorer, 26, 82, 881
Cosmic background radiation, 24, 26, 28, 30, 81-82, 297, 881-882, 963
Cosmic dust, 963
Cosmic microwave background radiation, 963. *See also* Cosmic background radiation

Cosmic Origins Spectrograph, 951
Cosmic rays, 73-78, 161, 232, 303, 660, 728, 745, 947, 963; cancer risks, 298; lunar impact, 391; solar wind and, 297
Cosmic scale factor, 871
Cosmic strings, 964. *See also* String theory
Cosmic X-Ray Explorer, 546-547
Cosmogony, 964
Cosmological constant, 24, 29, 800, 880
Cosmology, 78-84, 867-874, 964
COSTAR. *See* Corrective Optics Space Telescope Axial Replacement
Coulomb barrier, 553, 670
Coulomb repulsion, 550-551, 964
Couplings, 601-605
Covariant, 964
Cowan, Clyde L., 538
CP 1919 (pulsar), 653
Crab nebula, 953, 964; supernova, 817
Crab nebula pulsar, 257, 653-656, 662, 819
CRAF. *See* Comet Rendezvous and Asteroid Flyby
Crater morphology, 964
Craters, 964; Callisto, 347; Chicxulub, 512; cometary, 55; Earth, 497; Europa, 346; Ganymede, 246, 346; impact, 279-286, 421, 971; Jovian satellites, 214; lunar, 360-368, 378; Mars, 421-427, 631; Mercury, 461; meteor, 484; pedestal, 422, 979; Phobos and Deimos, 440; rampart, 422, 983; secondary, 984; Triton, 859; Venus, 921-925; volcanic, 421, 423, 446
Cratons, 132
Creation, 29; Earth, 103-108
Creation science, 106
Cretaceous-Tertiary event, 11, 283, 511-512
Critical density value, 83
Critical points, 331; hydrogen, 683
Critical universe, 875
Crossover time, 869, 872
Crust (defined), 964
Crust (Earth's), 115, 129-139, 186, 375, 582
Crustal differentiation, 964
Cryogenic volcanism. *See* Ice volcanism
Cryovolcanism, 235
Crystallization, 964
Crystals, 964
Cubewanos, 86, 353
Cumulates, 964
Curie temperature, 151, 964
Curvature radiation, 198
Curved space, 796, 964
Cyanobacteria, 110, 117, 355
Cyclotron radiation, 198, 964
Cydonia (Martian plain), 403

D″ layer, 125
D-type asteroids, 9
Dactyl (asteroid), 11
Dana, James D., 119
Dark clouds, 964
Dark energy, 81, 83, 800, 870, 876-877, 964
Dark matter, 82, 251, 502, 504, 800, 881, 964
Dark spots; Jupiter, 55; Neptune, 313, 316, 518, 520-524, 569; Uranus, 888
Dating; Earth's age, 105; radiometric, 105
Davis, Marc, 514
Davis, Raymond, Jr., 539-541
Dawn mission, 41, 43, 88
Dawson, George, 3
Day; defined, 964; Earth, 179, 965; lunar, 975; sidereal, 178; solar, 178
De revolutionibus (Copernicus), 599
December solstice. *See* Solstices
Declination, 66, 149, 157, 964
Deep Impact mission, 51, 61-62, 353, 557
Deep Near Infrared Survey, 666
Deep Sea Drilling Project, 138
Deep-sea trenches, 626
Deep space, 964
Deep Space 1, 51, 60
Deep space probes, 964
Deforestation, 102
Degenerate electron gas, 964
Degenerate matter, 656, 662, 818, 943-945. *See also* Electrons
Degradation, 965
Deimos, 8, 437-441, 620, 706
Deliberate design, 83
Denning, W. F., 497
Density, 965; bulk, 586; critical, 799; gas giants, 772; measurements, 771; planetary, 524; Pluto-Charon, 643; Saturn, 683; Triton, 857
Density separation, 375
Deoxyribonucleic acid, 231, 356; origins, 357
De Pater, Imke, 533-534
Descartes, René, 254, 580, 724, 776
De Sitter, Willem, 216, 875, 880
Despina (moon), 536-537
Diabase, 470
Diablo Canyon, 497
Diamond code (Gamow), 25
Diamonds, 163
Diapirism, 859
Dicke, Robert H., 26, 82, 732, 760
Differential rotation, 744, 753, 813
Differentiation, 965; asteroids, 10; Callisto, 37; crustal, 964; Earth, 115, 117, 139-145, 174, 770; Mercury, 461; meteoritic,

486-488; Miranda, 507; Moon, 369; terrestrial planets, 526; Triton, 858; Venus, 932
Digital imaging, 559, 561
Dinosaurs, 11, 16, 283
Diogenites, 469, 471, 487
Dione (moon), 204, 603, 703, 705, 707-708
Dipole fields, 145, 154, 159, 743, 965
Dirac, Paul Adrien Maurice, 943
Dirty snowball model, 44, 57, 555, 790, 965
Discontinuities, 186-187, 189, 965
Discontinuity boundaries, 135-139
Disk instability mechanism, 668
Disney, Michael J., 653
Dissociation, 965
Diurnal (defined), 965
Divergent plate boundaries, 130
DNA. *See* Deoxyribonucleic acid
Donati's comet, 48
Doppler effect, 561, 759, 965
Doppler radar, 604
Doppler shifts, 222, 257, 300, 303, 661, 732, 874; auroras, 19; exoplanet detection, 224-225, 228. *See also* Blueshifting; Doppler effect; Redshifting
Draconic month, 193, 965
Drake, Frank, 236, 335, 710
Dumouchel, Father, 48
Dunham, Edward, 896
Durrington Walls circle, 2
Dust (interstellar), 301. *See also* Cosmic dust; Interplanetary dust
Dust devils (Mars), 417, 567
Dwarf planets, 43, 85-89; Ceres, 6-7, 9, 11, 15, 40-44, 85, 211; definition, 211; Eris, 208-213; Pluto, 211
Dynamics Explorer satellites, 19, 22
Dynamo effect, 332, 965; Earth, 121, 145, 151; Neptune, 529; Saturn, 684, 686; Sun, 744, 813; Venus, 916, 932. *See also* Geodynamo theories
Dysnomia, 208-213

$E = mc^2$, 27, 199, 249, 398, 739, 847, 869
E-type asteroids, 9
Eagle nebula, 651
Earth, 844; age, 103-108; atmosphere, 108-114, 117, 564-565, 679; chemical composition, 115-120; core, 115, 121-125, 145, 151, 159, 187, 374, 582, 979; crust, 115, 129-135, 186, 582; density, 374; differentiation, 139-145; formation of, 140; heat flow, 163, 583; interior, 147; internal pressure, 583; internal temperatures, 583; landforms, 171; magnetic field, 121, 124-126, 145-158, 189; magnetopause, 159; magnetosphere, 156, 158-162, 590; mantle, 115, 121-122, 124, 126, 129, 140, 162-166, 186-188, 582, 586, 975; oblateness, 183; oceans, 58, 166-172; orbit around Sun, 96, 178; origins, 173-177; plate tectonics, 625; rotation, 95, 177-183, 612; shape, 178, 183-185; structure, 186-190; surface, 639; temperatures, 95; as terrestrial planet, 638, 843-847; thermal history, 115; tidal effects, 585; valleys, 441; wobble, 128
Earth day, 965
Earth-emitted radiation, 965
Earth inductors, 152
Earth-Moon system, 90-94, 179, 183, 368, 769; spin-orbit resonance, 602
Earth observation satellites, 102
Earth-orbital probe, 965
Earth-Sun relations, 94-99, 183, 399, 721
Earth system science, 99-103
Earth tide, 965
Earthquakes, 100, 102, 127-128, 134, 163, 165
Eccentricity, 93, 595, 965
Eclipse radiometry, 247
Eclipses, 190-197, 965; annular, 958; Jovian satellites, 195; lunar, 191, 979, 989; solar, 191-192, 196, 714, 716, 979, 989
Eclipsing binary detection method, 229
Ecliptic, 966, 981
Ecliptic coordinate system, 67
Ecliptic latitude, 67
Ecliptic longitude, 67
Ecliptic plane, 40, 96, 177, 180, 191, 203, 295, 352, 555, 576, 966
Eclogite, 163
Ecology, 966
Eddington, Sir Arthur Stanley, 273, 553, 726, 809, 942, 944
Eddy, John, 4, 753, 785
Edgeworth, Kenneth, 352
Edison, Thomas Alva, 663
Edlen, Bengt, 718
EEGs. *See* Evaporative gaseous globules
Effective temperature, 260, 966
Effelsberg radio telescope, 828
EGRET. *See* Energetic Gamma Ray Experiment Telescope
Einstein, Albert, 24, 29, 196, 248, 257, 732, 760, 797-798, 804, 847, 851, 865, 880-881, 944

Einstein crosses, 966
Einstein Observatory, 23
Einstein rings, 966
Einstein X-Ray Observatory, 546-547
Ejecta, 361, 966
Ejecta blanket, 370
Electric universe theory, 519
Electrojets, 20
Electrolysis, 966
Electromagnetic radiation, 73, 544, 547, 558, 561, 832, 964, 966, 982; early universe, 80, 869; nonthermal emissions, 197-200, 660; solar, 260; thermal emissions, 201-204, 660. *See also* Cosmic rays; Gamma rays; Infrared radiation; Microwaves; Radio waves; Spectroscopy; Ultraviolet radiation; Visible light; X rays
Electromagnetic spectrum, 544, 751, 966
Electron beam analysis, 170
Electron capture, 106
Electron volt, 966
Electrons, 966; big bang, 27; cosmic rays, 73; degenerate, 818, 943-945; electrojects, 20; magnetic fields, 17, 591; nucleosynthesis, 28; solar wind, 156; spectroscopy, 197
Electrophoresis, 966
Elementary particles, 966
Elements (chemical). *See* Chemical elements
Elliot, James, 896
Ellipses, 595, 966
Elliptical orbits, 595, 967
Elsasser, Walter, 151
Elysium region (Mars), 447
Emission lines, 560
Emission nebulae, 300
Emission spectra, 197, 199, 201, 967
Emissions; nonthermal, 197-200, 660; solar, 736-739, 778-784, 792-795; solar radio, 756-759; thermal, 201-204, 660. *See also* Electromagnetic radiation; Radiation; Spectroscopy
Enantiomers, 355, 967
Enceladus (moon), 204-208, 603, 608, 701-710, 898; life possibilities, 235; tectonics, 629; tidal stresses, 586
Encke (comet), 499
Endogenic metamorphism, 282
Energetic Gamma Ray Experiment Telescope, 239, 952
Energy, 967; fossil fuels, 850; nuclear, 850; solar, 850
Enstatite, 470, 768
Envelope solitons, 758
Eolian erosion, 967
Ephemerides, 46, 967
Epicenter, 967
Epimetheus (moon), 603, 701-710
Equatorial coordinate system, 66
Equatorial orbit, 967
Equilibrium; convective, 963; hydrostatic, 971; radiative, 983
Equivalence principle, 248, 255, 797, 981
Erg, 967
Eris (dwarf planet), 87, 208-213, 353, 573, 647
Eros (asteroid), 10, 12-14, 41; NEAR mission, 14
Escape velocity, 564, 967
ESO. *See* European Southern Observatory
Eta Carinae (nebula), 822
Ethane, 854
Euclid, 796
Eucrites, 469, 471, 487
Europa (moon), 213-220, 236, 244, 306, 308-309, 314, 316, 345-351, 621, 633; atmosphere, 217; craters, 214; density, 218, 579; ice flows, 216; magnetic field, 148; possible life, 219, 235; revolution, 350; surface, 215, 346; tectonics, 628; terrains, 215; tidal stresses, 586; water, 218, 235, 348, 350
European Southern Observatory, 580, 826
EUVE. *See* Extreme Ultraviolet Explorer
Evaporative gaseous globules, 651
Evection, 94
Event (defined), 967
Event horizon, 798, 967
Ewen, Harold, 303, 664
Exobiology, 231-237, 967. *See also* Astrobiology
Exoplanets, 227-231
EXOSAT, 951
Exospheres, 564, 967; Earth, 565
Expanding universe, 24, 78, 81, 251, 562, 799, 871, 874-879, 881
Explorer 1, 19, 22, 156, 158, 161, 912
Explorer 4, 912
Explorer 10, 789
Extinctions. *See* Mass extinctions
Extrasolar planets, 220-227, 292, 333, 578, 651, 777; detection, 227-231; life possibilities, 233; observation methods, 579-580
Extraterrestrial intelligence. *See* Search for extraterrestrial intelligence
Extraterrestrial life, 231-237, 357, 399
Extreme Ultraviolet Explorer, 833, 840, 864
Extremophiles, 406
Extrusive igneous rocks, 130

Faculae (solar), 715, 749, 785, 967
Failed stars. *See* Brown dwarf stars
False-color imagery, 967
Far Ultraviolet Spectroscopic Explorer, 836, 864
Faraday cup, 967
Feldspar, 116, 768, 770
Fermentation, 110
Fermi, Enrico, 943
Fernández, Julio, 87, 352
Ferraro, V. C. A., 788
Ferrous iron oxide, 769
51 Pegasi, 222, 229
55 Cancri, 222
Fireballs, 495, 497, 499, 511, 967
Fischer, Debra, 222
Fission (atomic), 967
Fission tracks, 967
Flammarion, Camille, 535
Flares. *See* Solar flares
Flash spectrum, 715
Flat space, 968
Flat universe, 875
Fleischer, Robert, 75
Fletcher, Leigh, 678
Flight paths, 968
Flow drapes, 448
Fluid mechanics, 968
Fluorescence, 968
Flux-gate magnetometers, 152
Flux ropes, 590
Flybys, 968
Flying sandbank model, 44
Focal ratio (f-number), 968
Foci, 968
Folland, D. F., 664
Forbidden lines, 71
Forces of nature, 26, 868, 877, 968
40 Eridani B, 942, 944
Fossil fuels, 263
Fossils, 355; anaerobic organisms, 110
Foucault, Jean-Bernard-Léon, 181
Foukal, Peter, 753
Fowler, Ralph Howard, 943
Fowler, William A., 77, 553, 809
Frail, Dale, 221, 228
Frame of reference, 968
Franklin, Kenneth, 316, 335
Fraunhofer, Joseph von, 561
Fraunhofer lines, 968. *See also* Absorption lines
Free return trajectory, 968
Frequency, 968
Friedman, Herbert, 546-547, 730, 794
Friedmann, Alexander, 24, 78, 871, 874, 880
Fumaroles, 168
Fundamental forces, 26, 868, 877, 968
FUSE. *See* Far Ultraviolet Spectroscopic Explorer
Fusion reactions, 398, 539, 549-554, 806, 968; protostars, 670; red dwarfs, 666; solar, 763-767; stellar, 847-852. *See also* Nucleosynthesis
Fusion reactors, 850

Gabbro, 770, 968
Gabrielle. *See* Dysnomia
Gain, 968
Galactic clusters. *See* Clusters and superclusters
Galactic coordinate system, 67
Galactic latitude, 67
Galactic longitude, 67
Galactic nuclei, 546
Galactic tides, 60, 62
Galatea (moon), 536-537
Galaxies, 28, 60, 874, 879-886, 968; active, 957; ages of, 30; clusters and superclusters, 26, 29, 78, 81, 292, 301, 504; formation, 81; interstellar matter in, 301; Milky Way, 58, 501-506, 546, 548, 977; origins, 504; Seyfert, 985; types, 504
Galilean moons, 345; Europa, 213-220
Galilei, Galileo. *See* Galileo Galilei
Galileo Galilei, 177, 213, 244, 305, 314, 335, 345, 561, 572, 599-600, 605, 616, 620-621, 623-624, 674, 698, 814
Galileo mission; asteroids, 11-12, 14; Europa, 214, 216, 219; Ganymede, 246; Io, 305, 307, 346; Jupiter, 217, 235, 315, 321, 327, 338-339, 341, 591
Galle, Johann Gottfried, 85, 528, 531, 624, 647
Galle ring (Neptune), 531
Gamma-ray astronomy, 238-244, 947-954, 968, 991
Gamma-ray bursters, 238-244, 793
Gamma-Ray Large Area Space Telescope, 837, 951
Gamma Ray Observatory, 239
Gamma rays, 544-545, 548, 739, 848, 862, 947, 968; supernovae, 820
Gamow, George, 25, 29, 79, 553
Ganymede (moon), 218, 244-248, 306, 309, 314, 345-351, 701; craters, 214; density, 218, 579; Jupiter's magnetosphere, 338; magnetic field, 148, 218-219; revolution, 350; surface, 346, 348; water, 244, 349-350
Gaps, 968
Garrels, Robert M., 169
Gas Chromatograph Mass Spectrometry Experiment (Viking mission), 234
Gas Exchange Experiment (Viking mission), 234
Gas giants, 230, 311-317, 333, 579, 774. *See also* Jovian planets; Jupiter; Neptune; Saturn; Uranus

Gaspra (asteroid), 11
Gatewood, George, 222
Gauss, Carl Friedrich, 6, 40, 152, 236, 798
Gegenschein, 968
Geiger counter, 969
Geller, Margaret, 884
Gellibrand, Henry, 150
Geminga (neutron star), 951
General relativity, 24, 29, 78, 82, 196, 248-252, 257, 562, 797, 803, 880, 944, 969, 984
General theory of relativity. *See* General relativity; Relativity
Geocentric orbits, 969
Geochemical sinks, 169, 969
Geochronology, 103-108, 969
Geodesy, 253, 969; solar, 732-736
Geodynamo theories, 152, 154-155, 159, 187, 969. *See also* Dynamo effect
Geomagnetic elements, 969
Geomagnetic storms, 72, 160
Geomagnetism, 969
Geometry, 969; Euclidean, 796; non-Euclidean, 796
Geomorphology, 969
Geostationary orbit, 969
Geosynchronous orbit, 969
Geothermal (defined), 969
Geotherms, 147
Geysers; Earth, 168; Enceladus, 235; Europa, 215-216, 218
GHGs. *See* Greenhouse gases
Ghost craters, 347
Giacconi, Riccardo, 546-547
Giant Magellan Telescope, 827
Giant molecular clouds, 300
Giant stars, 270
Gilbert, William, 149-150, 152
Gill, Sir David, 48
Giotto mission, 44, 50, 57-58
Glass, 969
Glass (lunar), 386, 391
GLAST. *See* Gamma-Ray Large Area Space Telescope
Gliese (brown dwarf), 32
Global circuit, 593
Global Oscillations Network Group, 735, 762
Global Surveyor. *See* Mars Global Surveyor
Global warming, 260-265, 420, 786; Venus, 936
Globular clusters, 546, 969
Gneiss, 117
Gold, Thomas, 29, 384, 655, 875
Gorchakov, E. V., 910
Gossamer Ring (Jupiter), 341-342
Granat observatory, 239, 241, 833
Grand maximum. *See* Medieval grand maximum
Grand Tour, 528, 887, 893
Grand unified theories, 29, 82, 877. *See also* Theories of everything; Unified field theories
Granite, 770, 969; Venus, 927-928, 933
Granules (solar), 715, 749, 764. *See also* Supergranules
Granulites, 133
Gravimeters, 255
Gravitation, 969
Gravitational acceleration, 185, 248, 253
Gravitational collapse, 576
Gravitational constant, 970
Gravitational contraction, 773, 970
Gravitational coupling, 350
Gravitational differentiation, 970
Gravitational fields, 253, 970; Earth, 256
Gravitational force, 970
Gravitational lensing, 250, 562, 797, 970
Gravitational mass, 970
Gravitational microlensing detection method, 228
Gravitational potential energy, 970
Gravitational tidal force, 91
Gravitational time dilation, 257
Gravitational wave detectors, 255
Gravitational waves, 257
Gravity, 969; measurement, 252-259; Newton's laws, 248, 253; orbital motion and, 594; relativity theory, 248-252; Sun, 734
Gravity assist, 970
Gravity gradiometers, 255
Gravity pendulums, 254
Grazing incidence, 970
Grazing-incidence telescopes, 947
GRB 050904, 242
GRB 080319B, 242
Great Bombardment, 921
Great Comet of 1843, 48
Great Daylight Comet of 1910, 49
Great Dying, 11, 284, 515
Great Observatories Program, 834
Great Red Spot (Jupiter), 55, 313-314, 316, 320, 324-330, 339, 522, 568, 614, 633-634
Great Rift (nebula), 503
Great September Comet of 1882, 48
Great White Spot (Saturn), 677
Green Bank Telescope, 829
Greenhouse effect, 98, 260-265, 970; Mars, 443; primordial, 171; Titan, 855-856; Triton, 859; Venus, 565, 638, 915, 919, 922, 936, 938, 941
Greenhouse gases, 260
Greenland, 119
Grigg-Skjellerup (comet), 51, 61
Grimaldi, P., 360

Grotrian, Walter, 717
Groundwater, 970
Guest star, 970
Gula Mons (Venus), 939
Gursky, Herbert, 547
Gusev crater (Mars), 410, 412
Gutenberg, Beno, 163, 587
Guth, Alan, 29
GUTs. *See* Grand unified theories
Gyroscopes, 970

H I and II regions, 300, 302-303, 503, 650, 970
Haar, Dirk ter, 724
Habitable zones, 222, 224, 230, 266-269, 667, 711, 928
Hadley cells (Venus), 566
Hadley Rille (Moon), 379
Hadrons, 970
Hale, George Ellery, 730, 743, 809, 811
Hale telescope, 825
Halemaumau Pit, Hawaii, 168
Halimede (moon), 536
Hall, Asaph, 437-438
Halley, Edmond, 46, 92, 150
Halley's comet, 44-52, 56-58, 62, 497; organic molecules, 233
Halo Ring (Jupiter), 342
Hammel, Heidi, 313, 518, 521, 888
Harding, Karl Ludwig, 465
HARPS. *See* High Accuracy Radial Velocity Planet Searcher
Hartmann, Johannes Franz, 303
Hartzes, Artie, 224
Hawaiian islands, 164; 168
Hawking, Stephen W., 548
Hayabusa probe, 41
HD 114762 (star), 223
HD 189733b (star), 224
HEAOs. *See* High-Energy Astronomical Observatories
Heat budgets, 94-99, 260, 970
Heat death, 971
Heat flow; Earth, 583; small bodies, 584
Heaviside, Oliver, 663
Hektor (asteroid), 8, 10
Helene (moon), 603, 707
Heliocentric model, 572, 599, 616
Heliocentric orbits, 971
Heliopause, 296, 790, 971
Helioseismology, 732, 741, 759-763
Heliosheath, 296
Heliosphere, 296, 592, 971
Heliospheric physics, 763-767
Helium, 297, 302, 550, 767, 808, 821; discovery, 716; interstellar medium, 300; solar, 739
Helium flash, 671, 724
Helium fusion, 398, 550, 807
Hellas Planitia (Mars), 422
Helmholtz, Hermann von, 726, 851
Henderson, Thomas, 882
Heracleides, 177
Herman, Robert C., 26
Hermes (asteroid), 8
Herschel, Sir John Frederick, 48, 535, 857
Herschel, Sir William, 48, 293, 501, 605, 624, 647, 674, 866, 884, 886, 893, 896, 901, 906
Hertz (defined), 971
Hertz, Heinrich Rudolph, 257, 662
Hertzsprung, Ejnar, 269, 272-273, 666
Hertzsprung-Russell diagram, 269-274, 398, 971
HETE. *See* High Energy Transient Experiment
Heterogeneous accretion model, 142
Hewish, Antony, 653, 662
Hexahedrites, 481, 971
Hey, J. S., 664
Hidalgo (asteroid), 8
Hide, Raymond, 325
High Accuracy Radial Velocity Planet Searcher, 228
High-Earth orbits, 971
High-Energy Astronomical Observatories, 75, 722
High-Energy Astrophysical Observatories, 951-952
High-energy astrophysics, 794
High Energy Transient Experiment, 951
High-gain antennae, 971
Highlands, 971
Highly elliptical orbits, 967
HiLat satellite, 22
Hilda (asteroid), 7
Hill, Henry, 733, 760
Hipparchus, 68, 92
Hipparcos satellite, 884
Hoagland, Richard, 403
Hoba West meteorite, 484
Hobby-Ebberly telescope, 834
Hodgson, R., 730
Hoffmeister, C., 791
Homogeneity problem, 881, 971
Homogeneous accretion model, 140, 971
Hooke, Robert, 150, 325
Horizon, 971
Horizon coordinate system, 66
Horizontal branch, 971
Horsehead nebula, 503
Hot Jupiters, 222, 224, 228, 230, 233, 292, 557, 578, 580
Hot spots, 164
Hour circle, 971
Howardites, 469, 471
Hoyle, Sir Fred, 29-30, 44, 77, 79, 553, 809, 875
H-R diagram. *See* Hertzsprung-Russell diagram
HST. *See* Hubble Space Telescope
Hubbard, William B., 336
Hubble, Edwin Powell, 24, 79, 82, 562, 726, 838, 871, 874, 877, 881, 884

Hubble age, 875
Hubble constant, 83, 875, 971
Hubble Space Telescope, 289, 559, 580, 834, 838, 842; auroras, 21; Ceres, 42; Eagle nebula, 651; Orion nebula, 651; Uranus, 888, 900
Hubble's constant, 838
Hubble's law, 28, 82, 874, 877, 971
Huchra, John, 884
Humason, Milton, 24, 874, 877
Humboldt, Alexander von, 496
Hurricanes, 178
Hut, Piet, 514
Hutton, James, 104, 441
Huygens, Christiaan, 254, 314, 428, 605, 624, 698
Huygens probe, 235, 315, 634, 853
Hvatum, S., 335
Hydrocarbons, 971; comets, 557; Jovian, 339; Messier 101, 841; Titan, 569, 634, 702, 855; Triton, 858; Uranus, 886, 890
Hydrogen, 297, 302, 660, 664, 670, 767, 808, 821; interstellar medium, 300; liquid metallic, 311, 332, 683; solar, 715, 739. *See also* Liquid metallic hydrogen
Hydrogen bomb, 851
Hydrogen fusion, 398-400, 550, 806, 849
Hydrostatic equilibrium, 971
Hydrothermal vents, 168, 357
Hyperion (moon), 276, 701-710

Iapetus, 275-279, 603, 705, 707
IAU. *See* International Astronomical Union
Icarus (asteroid), 9-10
ICB. *See* Inner core boundary
ICE. *See* International Cometary Explorer
Ice, 772, 774; comets, 555; Europa, 216, 218; Jovian satellites, 345; liquid, 331; outer planets, 585; Saturn, 683; Titan, 853. *See also* Water
Ice caps. *See* Polar caps
Ice giants, 890, 892-893
Ice volcanism, 218, 585, 621
Ida (asteroid), 11
IEU. *See* International Ultraviolet Explorer
Igneous rocks, 129, 964, 971; lunar, 391; Venus, 927-928, 933
IGY. *See* International Geophysical Year
Imaging, 971
Imaging Compton Telescope, 239, 952
Imbrium Basin (Moon), 462
IMP satellites, 75
Impact basins, 971
Impact craters, 279-286, 627, 776, 921, 971; comets, 55; Earth, 484, 497, 556; Io, 305; lunar, 360-362, 365, 369, 378, 381, 391; Mars, 421, 631; Mercury, 460; Uranian satellites, 902; Venus, 921-925, 931, 939. *See also* Craters
Impactor hypothesis of lunar origin, 622
Inclination, 149, 157
Inclusions, 972
Index of refraction, 972
Inertia, 594, 972
Inertial frames of reference, 801
Inertial mass, 972
Inferior conjunction, 962
Inflation theory, 29, 972
Inflationary universe, 868
Infrared Astronomical Satellite, 288, 294, 579, 825, 838
Infrared astronomy, 286-295, 649, 651, 972
Infrared radiation, 286-295, 544, 972; solar, 736-739. *See also* Infrared astronomy
Infrared Space Observatory, 289, 836
Infrared spectrometer, 972
Infrared Telescope Facility, 293
Ingersoll, Andrew, 321, 339
Inhomogeneous accretion model, 142
Inhomogeneous accretion theory, 972
Inner core boundary, 121, 123, 189
INTEGRAL. *See* International Gamma-Ray Astrophysics Laboratory
Intelligent design, 83
Intelligent life, 405; Mars, 401
Intercloud gas, 299
Intercrater plain, 972
Interferometers, 829, 972
Interferometry, 972
Intergalactic medium, 972
International Astronomical Union, 87, 211, 573
International Cometary Explorer, 45
International Gamma-Ray Astrophysics Laboratory, 951
International Geophysical Year, 21, 730, 912, 972
International Space Station, 102, 393; micrometeoroid damage, 298
International Sun-Earth Explorer, 239
International Ultraviolet Explorer, 45, 722, 838, 863
International Union of Geodesy and Geophysics, 129
Interplanetary dust, 295, 297

Interplanetary environment, 295-299
Interplanetary magnetic field, 159
Interplanetary matter, 297
Interplanetary space probes, 964
Interstellar clouds, 299-304
Interstellar dust, 963
Interstellar gas, 302
Interstellar medium, 299-304, 972
Intertia, moment of. *See* Moment of inertia
Interval, 972
Intrusive igneous rocks, 130
Invariant, 972
Inverse beta decay, 538, 540, 973
Io (moon), 213, 218, 244, 304-310, 314, 345-351, 591, 621, 633; atmosphere, 306; chemistry, 306; composition, 347; density, 218, 579; flux tube, 591; heat flow, 586; interior, 587; Jupiter's magnetosphere, 338-339; mountains, 306; plasma torus, 307, 338, 591; plumes, 346; radio emissions, 347; revolution, 350; sulfur, 770; surface, 346; tectonics, 628; compared to Venus, 941; volcanism, 219, 293, 307, 347, 591
Ionization, 973
Ionosphere (Earth's), 18, 20-21, 565, 973; disruptions, 160
Ions, 17, 973
Iraki, G., 292
IRAS. *See* Infrared Astronomical Satellite
IRAS-Araki-Alcock (comet), 292
Iridium, 16, 283-284, 511, 513-514, 973
Iron; Earth, 124, 582; Mars, 584; meteorites, 483; Moon, 584; nucleosynthesis of, 544, 551, 671, 807, 849; solar wind, 789; supernovae, 818
Iron catastrophe, 121
Iron meteorites, 115, 143, 468. *See also* Nickel-iron meteorites
Iron-nickel metal, 768
Ishtar Terra (Venus), 631, 915, 930
Isidis (Mars), 422
Island arcs, 626
ISO. *See* Infrared Space Observatory
Isogonic maps, 150
Isomagnetic charts, 973
Isoporic charts, 150
Isostasy, 132, 973
Isotopes, 106, 973
Isotropic (defined), 973
Itokawa (asteroid), 41
IUE. *See* International Ultraviolet Explorer
Ivuna (meteorite), 473
Ixion (plutino), 88

Jackson, Albert, 515
James Webb Space Telescope, 290, 519, 726, 838
Jansky, Karl Guthe, 504, 659, 663
Janssen, Jules, 716
Janus (moon), 603, 701-710
Jeans, Sir James Hopwood, 581, 724
Jewitt, David, 85, 353
Jezero crater (Mars), 445, 458
JIMO. *See* Jupiter Icy Moons Orbiter
Jodrell Bank telescope, 828-829
Jovian planets, 311-317, 524, 579; atmospheres, 568; composition, 770. *See also* Jupiter; Neptune; Saturn; Uranus
Jovian satellites, 216; Jupiter's magnetosphere, 338
Jovian system. *See* Jovian planets; Jovian satellites; Jupiter
June solstice. *See* Solstices
Juno (asteroid), 7, 9; discovery, 40
Jupiter, 23, 33, 115, 213, 225, 311-317, 633; atmosphere, 112, 318-324, 564, 568, 632, 679; auroras, 21; comets and, 52-56, 352; Galileo mission, 235; Great Red Spot, 324-330, 568, 970; interior, 331-335, 586; Io, 304-310; Jovian system, 218; magnetic field, 148, 335-341; magnetosphere, 307-308, 591; oblateness, 312, 331; radiation, 335-341; radiation belts, 306; rings, 341-344, 605-611; rotation, 614; satellites, 195, 213, 304, 345-351, 621, 633. *See also* Jovian planets
Jupiter Icy Moons Orbiter, 216, 350
Jupiter-like planets, 220, 222-225, 973
JWST. *See* James Webb Space Telescope

K-band, 973
K phases, 127
Kaiser, Frederick, 437
Kamacite, 481, 488, 973
Kamiokande neutrino detectors, 540, 542
Kant, Immanuel, 92, 580, 618, 724, 776, 851, 874
Kapteyn, Jacobus Cornelis, 303
Karoo Array Telescope, 829
KBOs. *See* Kuiper Belt objects
Keck telescopes, 826, 834
Kelvin, Lord (Sir William Thomson), 726, 730, 851

Kelvin-Helmholtz contraction, 332, 684
Kelvins, 973
Kennelly, Arthur Edwin, 663
Kepler, Johannes, 48, 90, 177, 595, 597, 603, 616, 817
Kepler mission, 230
Kepler Observatory, 233, 580, 837
Kepler's laws of motion, 973
Kerr, Roy, 799
Kilogram, 973
Kilometer, 973
Kimberlite, 163, 586
Kinetic energy, 973
Kinetic pressure, 973
Kirkpatrick-Baez telescope, 948
Kirkwood, Daniel, 7, 497, 603, 698
Kirkwood gaps, 7, 602-604, 698
Kitt Peak telescope, 829
Kleinmann-Low nebula, 290
Kohoutek (comet), 57
Komatiites, 119
Korolev Crater (Mars), 431
Koronis asteroids, 11
Kowal, Charles, 85, 353
Krakatoa eruption, 513
KREEP-norite, 369-370, 973
Krypton, 789
K/T boundary, 513
K-T event. *See* Cretaceous-Tertiary event
Kuiper, Gerard Peter, 58, 218, 352, 581, 624, 901
Kuiper Airborne Observatory, 288, 307, 899
Kuiper Belt, 43, 58-59, 86, 208, 212, 295, 352-354, 510, 557, 774, 973
Kuiper Belt objects, 86, 352, 647
Kunzig, Robert, 936

L waves, 164
Labeled Release Experiment (Viking mission), 234
Lagrange, Joseph-Louis, 8, 602, 604
Lagrangian points, 290, 602, 721, 838, 973
Lakes; lava, 305; Mars, 411, 445, 458; Titan, 235, 634, 707, 853; Triton, 859
Lakshmi Planum, 939
Lalande, Joseph-Jerome de, 46
Lalande 21185 (red dwarf), 222
Langley, Samuel Pierpont, 293, 752, 784
Laomedeia (moon), 536
Laplace, Pierre-Simon, 216, 577, 580, 618, 724, 777
Lapse rate, 974
Large Aperture Mirror Array, 828
Large Binocular Telescope, 826
Large impact hypothesis, 368
Large Millimeter Telescope, 829
Larissa (moon), 536
Larson, Richard, 724
Laser guide stars, 826
Laser Interferometer Gravitational-Wave Observatory, 256
Laser Interferometer Space Antenna, 256
Laser radiation, 199
Laser ranging, 974
Laser reflectors, 974
Lasers, 974
Lassell, William, 535, 624, 857, 901
Lassell ring (Neptune), 532
Latham, David, 223
Latitude, 974
Laughlin, Greg, 666
Launch sites, 974
Lava, 974
Lava domes (Venus), 939
Lava flows; Io, 628; lunar, 378; Mars, 443, 447-448; Moon, 627; Venus, 628, 931, 938
Lavoisier, Antoine, 495
Leavitt, Henrietta Swan, 874, 883
Lehmann, Inge, 123, 587
Leighton, Robert, 732, 759
Lemaître, Georges, 24, 79, 871, 874, 877
Leonid meteor showers, 295, 496-497
Lepaute, Nicole, 46
Leptons, 27, 974
Le Verrier, Urbain 497, 527, 531
Le Verrier ring (Neptune), 532
Levy, David, 52
Lewis, John, 218, 319
Libration, 93, 601
Lick Observatory, 825; search for extraterrestrial intelligence, 712
Life; atmospheric conditions for, 420; definition, 405; Europa, 219; evolution of, 110; extraterrestrial, 222, 399, 636; habitable zones, 266-269; intelligent, 401; interstellar chemicals, 303; Mars, 405-409, 635; meteorites, 474, 493; organic molecules, 473-474, 476; origins, 110, 117, 355-360; potential on Jovian moons, 316; search for extraterrestrial intelligence, 710-713; Sun's influence, 94; supernovae and, 821; water, 171, 266
Light curve, 974
Light-years, 832, 974
Lightfoot, Joseph Barber, 104
Lightning, 156; Venus, 919, 940
LIGO. *See* Laser Interferometer Gravitational-Wave Observatory
Limb, 974
Limb darkening, 384, 737, 780

Limestone, 118
Liquid ices, 331, 683
Liquid metallic hydrogen, 148, 311, 332, 336, 338, 526, 586, 615, 675, 679, 683, 686, 974
LISA. *See* Laser Interferometer Space Antenna
Lithification, 974
Lithium, 974
LITHOPROBE program, 138
Lithospheres, 974; Earth, 119, 130, 135, 163, 165, 186; lunar, 375; Mercury, 461
Little Green Men theory, 653, 662, 712
Little Ice Age, 97, 750, 785
Little Maunder minimum, 745
Little Red Spot, 327, 329
Lobachevsky, Nikolai Ivanovich, 796, 798
Local Group of galaxies, 504, 874, 879
Local inertial frame, 801
Local thermodynamic equilibrium, 738
Lockwood, G. Wesley, 521
Lockyer, Norman, 716
Lodge, Sir Oliver Joseph, 663
Lodranites, 488
Loki (volcano), 305-307
Longitude, 974
Look angle, 974
LoPresto, James C., 753
Lorentz, Hendrik Antoon, 803
Lorentz force, 122
Lorentz transformations, 801, 803
Low, Frank, 294
Low-Earth orbit, 974
Lowell, Percival, 401, 428, 430
Lowell bands, 435, 975
Ludwig, George H., 912
Luminosity, 269, 975; stellar, 987
Luna missions, 363, 376, 382, 384, 386, 394, 491, 619, 789, 910, 912
Lunar. *See* Moon (Earth's)
Lunar day, 975
Lunar eclipse, 191
Lunar laser ranging, 181
Lunar month, 975
Lunar Orbiters, 363, 376
Lunar Prospector, 364, 372, 376, 382, 387, 619
Lunar Reconnaissance Orbiter, 372, 382, 387
Lunar samples, 378
Lunokhod rovers, 394
Luu, Jane, 85, 353
Luyten, Willem Jacob, 942
Lyell, Sir Charles, 104
Lyman Observatory, 864
Lyot, Bernard, 718
Lyrae stars, 984
Lyttleton, R. A., 44

M-branes, 83
M16 nebula, 651
M20 nebula, 725
M-type asteroids, 9
Mab (moon), 897
McCauley, J. F., 422
McCord, Thomas, 536
McDonald, Frank, 75
McDonald Observatory, 826, 834
Mach, 975
McIlwain, Carl E., 912
Mackenzie, Fred T., 169
MacKie, Euan, 1, 5
Mafic rocks, 975
Magellan mission, 628, 631, 918, 921, 929, 934, 940
Magellanic Clouds, 239, 820, 975; interstellar matter in, 301
Magma, 975
Magma chambers, 165
Magnetars, 293
Magnetic anomalies, 156, 975
Magnetic fields, 17-19, 124-125, 161, 296, 561, 589-594, 975; effect on cosmic rays, 73-74; Earth, 19-20, 121, 126, 145-149, 154-159, 189; Europa, 148, 218; Ganymede, 148, 218, 246, 346; interstellar, 301; Jovian planets, 313; Jupiter, 148, 311, 318, 332, 335-341, 633; magnetars, 293; Mars, 21, 147; Mercury, 148, 460, 463; Moon, 147, 376; Neptune, 525, 527-530; origins, 145-149; Saturn, 148, 311, 591, 675, 684, 686-692; secular variations, 126, 149-154; solar, 70, 148, 715, 718, 727, 741, 743-749, 753, 762, 792, 811-815; solar storms and, 297; solar wind, 790; stellar, 548, 655, 746, 814; Uranus, 893-896; Venus, 148, 916, 932
Magnetic knots, 745, 812
Magnetic poles, 146, 154, 975
Magnetic reversals, 150, 153, 156, 189, 284, 593; Earth, 16, 122, 124, 146; Sun, 744; Uranus, 894; Venus, 916
Magnetic storms, 71, 975
Magnetic surveys, 152, 975
Magnetite, 116
Magnetograms, 749
Magnetographs, 746
Magnetohydrodynamics, 23, 332, 721
Magnetometers, 157, 160, 746, 975; flux-gate, 152; lunar, 396; proton precession, 152; rubidium vapor, 152; superconducting, 152
Magnetopauses, 590, 790, 975; Earth, 156, 159-160; Jupiter, 337-338; Saturn, 689; Uranus, 894
Magnetosheaths, 590, 975; Earth, 160, 910; Jupiter, 338
Magnetospheres, 589-594, 975; Earth, 19-20, 158-162,

316, 590, 790, 910; Jovian, 309; Jupiter, 312, 335-339, 591; Mars, 591; Mercury, 590; Neptune, 592; Saturn, 312, 591, 687-688; Uranus, 592, 894; Venus, 590
Magnetotails, 589-590, 975; Earth, 159, 590; Jupiter, 337-338, 591; Uranus, 894
Magnitude, 975
Main Ring (Jupiter), 341-342
Main sequence stars, 270, 398-401, 544, 670, 806, 818, 975; Sun, 577, 723
Makemake (dwarf planet), 212
Mantles; Earth, 115, 121-122, 124, 126, 129, 135-140, 162-166, 186-188, 582, 586; Neptune, 525; Uranus, 890, 894
Maraldi, Giacomo, 428
Marconi, Guglielmo, 663
Marcy, Geoffrey, 222-223, 225, 230
Maria, 975
Mariana Trench, 131, 626
Mariner missions, 416, 429, 431, 449, 628; Mariner 1, 916; Mariner 2, 630, 789, 916; Mariner 4, 402, 418, 424, 444, 630, 640; Mariner 5, 917; Mariner 6, 418, 444; Mariner 7, 418, 444; Mariner 9, 402, 430, 437, 443-444, 448, 452, 640; Mariner 10, 462, 590, 630, 634, 826, 917
Maritime satellites, 975
Marius, Simon, 213
Marius Hills (Moon), 379
Mars, 115, 234, 597, 607, 634, 845; atmosphere, 108, 111-112, 415-421, 442, 564, 566; auroras, 21; canals, 401; carbon dioxide, 416, 442; channels, 401, 428, 441, 454; craters, 421-427; face, 403; illusions and conspiracies, 401-405; interior, 584; life possibilities, 232, 235, 405-409; magnetic field, 147; magnetosphere, 591; meteroids, 490-494; polar caps, 401, 416, 428-437, 453, 632, 635, 639-640; precipitation, 430, 442; rotation, 614; satellites, 437-441, 621; surface, 639; surface experiments, 409-415; tectonics, 585, 628; temperatures, 442; as terrestrial planet, 639, 843-847; valleys, 441-446; volcanoes, 446-452; water, 452-459; winds, 417, 443
Mars Climate Orbiter, 432
Mars Exploration Rovers, 235, 407, 410, 412-413, 455, 567, 591, 631
Mars Express, 21, 433
Mars Global Surveyor, 21, 432, 443
Mars Observer, 432
Mars Odyssey, 433
Mars Pathfinder, 235, 409, 412, 432, 452, 454, 630-631
Mars Phoenix, 413, 425, 434, 444, 456, 632, 640
Mars Polar Lander, 433, 456, 632
Mars Reconnaissance Orbiter, 434, 445, 457
Martians, 401
Mascons, 370, 381, 976
Masers, 976
Mass, 976
Mass-energy equivalence, 27, 199, 249, 398, 739, 847, 851, 869
Mass extinctions, 11, 16, 55, 283-284, 366, 500, 510, 514, 556, 598
Mass spectrometer, 976
Mass spectrometry, 169
Masursky (asteroid), 14
Matese, John, 515
Mathilde (asteroid), 14
Matrix, 976
Matter, 976; creation of, 27, 869
Mauna Kea observatory (Hawaii), 287
Maunder, E. Walter, 785
Maunder minimum, 97, 593, 715, 745, 785, 976
Maxima. *See* Medieval grand maximum
Maxwell, James Clerk, 257, 605, 624, 662
Maxwell Montes (Venus), 631, 916, 930
Maxwell's equations, 803
Mayer, J. Robert, 725
Mayor, Michel, 222, 229
Mazomba Crater (Triton), 859
Medicine wheels, 3
Medieval grand maximum, 98, 716, 750, 785
MeerKat telescope, 829
Megaparsec, 976
MER. *See* Mars Exploration Rovers
Mercury, 115, 213-214, 267, 460-468, 630, 634, 843; atmosphere, 111, 637, 825; craters, 421, 627; formation, 769; lack of atmosphere, 564; magnetic field, 148; magnetosphere, 590; rotation, 96, 613; solar orbit, 250, 257, 798; spin-orbit coupling, 602, 613; temperatures, 96; as terrestrial planet, 843-847; tidal locking, 604
Meridiani Planum (Mars), 411
Mesosiderites, 487, 489
Mesosphere (Earth's), 565, 976
MESSENGER mission, 464
Messier, Charles, 976
Messier numbers, 976
Metabolic process, 976
Metallic hydrogen. *See* Liquid metallic hydrogen
Metamorphic rocks, 129-130; lunar, 391

Metamorphism, 976; endogenic, 282; shock, 282
Meteor Crater, 484
Meteor showers, 295, 495-501, 976
Meteorites, 73, 115, 143, 578, 976; achondrites, 468-472; ages, 119, 476; ataxites, 959; carbonaceous chondrites, 961; chemistry, 578; chondrites, 9, 167, 472-480, 578, 583, 769, 962; composition, 771; formation, 776; hexahedrites, 971; history, 495-501; iron tools, 483; lunar impacts, 360, 368; lunar rocks, 90; microfossils, 493; Murchison, 233-234; nickel-irons, 481-486; organic compounds in, 355; planet formation, 578; sources, 9; stony irons, 9, 486-490; water, 167. *See also* Achondrites; Carbonaceous chondrites; Chondrites; Meteor showers; Nickel-iron meteorites; SNC meteorites; Stony iron meteorites
Meteoritics, 976
Meteoroids, 295, 976; damage from, 298; from Mars, 490-494; from the Moon, 490-494
Meteors, 976. *See also* Asteroids; Comets; Meteor showers; Meteorites
Meter, 976
Methane; atmospheric, 261, 263; Charon, 644; cometary, 46; greenhouse gas, 264; Jupiter, 769; Neptune, 518; planetesimals, 774; Pluto, 643; Titan, 852-855; Triton, 859; Uranus, 886
Metis (moon), 341
Metric system, 976
Metric ton, 976
MGS. *See* Mars Global Surveyor
Mica, 116
Michel, Helen V., 514
Microfossils, 358
Microgravity, 976
Micrometeorites, 977
Micrometeoroids, 295, 499, 977; damage from, 298
Micron, 977
Micropaleontology, 977
Microwaves, 544, 977
Mid-Atlantic Ridge, 131, 137, 626
Milanković, Milutin, 97
Milky Way, 58, 104, 303, 501-506, 552, 773, 864, 870, 879, 884, 953, 977; infrared sources, 288; interstellar matter in, 301; interstellar medium, 302; supernovae, 816, 822
Miller, Stanley, 231
Miller-Urey experiment, 231, 233
Miller-Urey synthesis, 232, 977
Millibar, 977
Milne, John, 587
Mimas (moon), 603, 608, 688, 698, 703-704, 708
Minerals (lunar), 391
Mini black holes, 251
Minima. *See* Maunder minimum; Spörer minimum
Mink, Douglas, 897
Minkowski, Hermann, 804
Miranda (moon), 21, 506-509, 579, 621, 624, 644, 902
Missing mass, 964
MOA-2007-BLG-192LB (exoplanet), 224
Model, 977
Moho. *See* Mohorovičić discontinuity
Mohole project, 138, 163
Mohorovičić, Andrija, 163, 587
Mohorovičić discontinuity, 130, 133, 135-139, 162-163, 186, 189
Molecular clouds, 300, 650, 805, 977
Molecular dissociation, 453
Molecules, 977
Moment of inertia, 586
Month; synodic, 975
Month (terrestrial), 90
Monuments of Mars, The (Hoagland), 403
Moon (Earth's), 90-94, 213-214, 236, 305, 587, 619, 769; absence of water, 234; albedo, 460; atmosphere, 619; colonization, 382, 623; composition, 380-381, 389; core, 391; craters, 360-368, 421, 627; density, 374; eclipse, 191, 194; erosion, 391; highlands, 366, 370, 380, 390, 460; history, 368-374, 379; impact craters, 279-280; infrared radiation, 286; interior, 374-378, 584; lack of water, 170; magnetic field, 147; maria, 370, 378-384, 390; meteroids, 490-494; mining potential, 623; orbit around Earth, 179; origin, 92, 368, 774; rocks, 389-394; soil samples, 384-389, 396; spin-orbit resonance, 602; surface, 366, 370, 379; surface experiments, 394-397; synchronous rotation, 620; temperatures, 95-96; tidal effects, 585; umbra, 192; valleys, 441
Moon rocks. *See* Lunar samples; Moon (Earth's)
Moonquakes, 375, 585
Moons, 977. *See also* Moon (Earth's); Planetary satellites; Satellites; *specific names*
Morehouse's comet, 48
Morgan, William Wilson, 303

Motion; Kepler's laws, 595-596, 616, 973; Newton's laws, 594, 597-598, 616
Motion, laws of; planetary, 597
Moulton, Forest, 227, 581
Mount Wilson Observatory, 874
Mountains, 132; Ganymede, 346; Io, 306, 345; Mars, 639; Mercury, 462; Moon, 370; Venus, 930
Mr. Tompkins series (Gamow), 25
Muller, Richard, 284, 514
Multidimensional membranes, 83
Murchison (meteorite), 233-234, 474
Murcray, David A., 754

Naiad (moon), 536
Nakhlites, 470
Namouni, Fathi, 533
Nanometer, 977
Navigation at sea, 153
Near-Earth space, 977
Near Infrared Camera and Multi-Object Spectrometer, 289, 533
Near-infrared rays, 977
NEAR mission, 12, 14, 41
Nebulae, 81, 398, 805, 818, 977, 982; Coalsack, 503; Crab, 964; emission, 300, 967; Great Rift, 503; Horsehead, 503; primordial solar, 981; reflection, 301, 983; solar, 986; stellar, 670
Nebular condensation, 906
Nebular hypothesis, 41, 578, 580, 618, 649, 724, 773, 776, 977
Nemesis theory, 55, 284, 510-516
Neptune, 115, 311-317, 634, 647, 892; atmosphere, 112, 516-520, 564, 568; auroras, 21; discovery, 85, 889; Great Dark Spots, 520-524, 569; interior, 524-527, 529, 586; magnetic field, 527-530; magnetosphere, 592; perturbation, 572; rings, 530-535, 537, 605-611; rotation, 615; satellites, 535-538, 621-622, 624. *See also* Jovian planets
Neptune Orbiter Probe, 519, 523, 529
Nereid (moon), 535, 621-622, 647
Neso (moon), 536
Neumann bands, 977
Neutral gas analyzer, 977
Neutrino astronomy, 538-543
Neutrino detectors, 538
Neutrinos, 27, 297, 400, 550, 741, 816, 819-820, 848, 850, 977
Neutron degeneracy, 819
Neutron stars, 221, 240-241, 272, 539, 544-546, 662, 807, 809, 819, 850, 977; pulsars as, 653-658
Neutrons, 27, 106, 977
New astronomy, 293, 978
New General Catalog of Nebulae and Clusters of Stars (Dreyer), 978
New Horizons mission, 43, 212, 246-247, 321, 327, 354, 634, 642, 646; Jupiter, 341
New Technology Telescope, 826
Newgrange, 2
Newton (unit), 978
Newton, H. A., 497
Newton, Sir Isaac, 24, 46, 92, 177, 194, 253, 254, 335, 561, 580, 594, 599-600, 603, 616, 776, 832
Newton's laws, 253. *See also* Gravity; Motion
Nickel-iron meteorites, 468, 481-486. *See also* Iron meteorites
NICMOS. *See* Near Infrared Camera and Multi-Object Spectrometer
Niedner, Malcolm B., 791
Nimbus 7, 785
1989 N1 (Neptunian satellite), 621
Nitrogen; atmospheric, 109; Triton, 859
Nitrous oxide; atmospheric, 261; greenhouse gas, 264
Noble gases, 789
Nonthermal emissions, 197-200, 660; solar, 728
Nördlinger Ries crater, 280
Nordman, Charles, 663
Normal-incidence telescopes, 948
North magnetic pole, 155
North point, 978
North Star, 180
Northwest Territories, 119
Novae, 543-549, 816-823, 945, 970, 978
Novikov, Igor, 548
Noyes, Robert, 224
Nuclear burning, 978
Nuclear energy, 850, 978
Nuclear reactions, 978. *See also* Fusion reactions; Nucleosynthesis
Nucleic acids, 231
Nucleogenesis, 869. *See also* Nucleosynthesis
Nucleosynthesis, 25, 73, 75, 81, 297, 301-303, 544, 549-554, 657, 672, 767, 773, 807, 818, 820-821, 869, 953, 978; solar, 739-742. *See also* Chemical elements; Fusion reaction
Nucleotides, 356
Nucleus (defined), 978
Nutation, 616; Earth, 180
Nyquist, L. E., 790

OAOs. *See* Orbiting Astronomical Observatories
Oberon (moon), 903
Oberth, Hermann, 833
Oblate spheroid, 978
Oblateness, 312; Sun, 732-733

Occultations, 193, 645, 978
Ocean crust, 131, 136
Ocean currents, 178
Ocean Drilling Project, 138
Ocean of Storms (Moon), 370
Ocean ridge-rift system, 131
Ocean surface, 184
Oceanic crust, 142, 186
Oceans, 166-172; cometary origins, 58; Earth, 95; origin, 109, 166; saltiness, 169
Octahedrites, 481, 978
Odyssey. *See* Mars Odyssey
Odyssey mission, 640
OGLE. *See* Optical Gravitational Lensing Experiment
Olbers, Heinrich Wilhelm, 7, 40, 497
Oldham, Richard Dixon, 123, 587
Olivine, 116, 187, 769, 978
Olmstead, D., 497
Olympus Mons (Mars), 402, 423, 448, 628, 631
1RXS J160929.1-210524 b, 226
Oort, Jan Hendrik, 55, 58, 555
Oort Cloud, 55, 58-59, 61-62, 284, 352, 510, 555-558, 774, 978; Pluto and, 647
Open universe, 875, 978
Ophelia (moon), 899
Ophiolites, 136-137, 586
Ophiuchus (constellation), 227
Opik, Ernst, 553
Oppenheimer, J. Robert, 656
Oppenheimer-Volkoff limit, 819
Opportunity (Mars Rover), 235, 411, 413. *See also* Mars Exploration Rovers
Opposition, 978
Optical astronomy, 558-563. *See also* Telescopes
Optical Gravitational Lensing Experiment, 230
Optical radiation, 978
Optical window, 978
Orange giant stars, 270, 672
Orbital couplings, 601-605
Orbital eccentricity, 979
Orbital inclination, 979
Orbital parameters, 979
Orbital periods, 979-980
Orbital phase detection method, 229
Orbital resonance, 601-605
Orbiting Astronomical Observatories, 833, 838, 863. *See also* Copernicus
Orbiting Solar Observatories, 719, 779, 833, 951-952
Orbits, 594-601, 777, 978; circular, 962; elliptical, 967; equatorial, 967; geocentric, 969; geostationary, 969; geosynchronous, 969; heliocentric, 971; high-Earth, 971; parabolic, 979; parking, 979; polar, 981; prograde, 981; retrograde, 984; transfer, 989
Orcus (plutino), 88
Ordinary chondrites, 979
Organic molecules, 231, 355, 979
Orgueil (meteorite), 473
Oriented Scintillation Spectrometer Experiment, 239, 952
Orion Crew Exploration Vehicle, 372
Orion nebula, 649
Orionid meteor showers, 497
Orogenic belts, 132
Orton, Glenn, 519
Oscillating universe, 979
Oscillations. *See* Solar oscillations
Osiris (exoplanet), 224
OSOs. *See* Orbiting Solar Observatories
OSSE. *See* Oriented Scintillation Spectrometer Experiment
Ostwald, Friedrich Wilhelm, 272
Outer space, 979
Outgassing, 453, 564, 979; comets, 62; Earth, 108, 167-168
Overwhelmingly Large Telescope, 828
Oxygen, 772; atmospheric, 110, 565
Ozone, 261, 566, 782, 979; atmospheric, 110, 261, 263; greenhouse gas, 264
Ozone layer, 565, 782

P modes, 734
P waves, 123, 125, 127, 132-133, 143-144, 163-164, 188, 981; shadow zone, 189
Pacini, Franco, 656
Pair annihilation, 199
Pair production, 869
Paleomagnetism, 152, 979
Palitzsh, Johann Georg, 46
Pallas (asteroid), 7
Pallasites, 486, 488-489
Palomar Observatory, 825
Pancake domes (Venus), 939
Pandora (moon), 707
Panspermia, 979
Parabolic orbits, 979
Parallax, 979
Parallax method, 882
Parallax shifts, 513
Parisi, Mirta, 907
Parker, Eugene N., 718, 744, 787, 789, 813
Parking orbits, 979
Parsec, 979
Partial melting, 979
Partial solar eclipse, 193
Particle horizon, 971
Pathfinder. *See* Mars Pathfinder
Pathological science, 404
Patroclus (asteroid), 8
Pauli, Wolfgang, 538
Pauli exclusion principle, 979
Pavonis Mons (Mars), 423, 448

PcP waves, 127
PcS waves, 127
Peale, Stanton, 587
Pearson, Mike Parker, 2
Pedestal craters, 422
Peebles, P. J. E., 26
Pele (volcano), 306-307
Pellas (asteroid), 40
Pendulum clocks, 181
Penston, Michael, 724
Penumbra, 980; sunspots, 812
Penumbral lunar eclipse, 191
Penzias, Arno A., 26, 82, 871, 875
Perfect spheroid, 980
Periapsis, 980
Pericynthion, 980
Peridotite, 143, 163, 186, 770, 980
Perigee, 980
Perihelion, 96, 178, 595, 980
Perilune, 980
Period, 980
Period-luminosity scale, 883
Permafrost, 980; Mars, 112, 406, 417, 422, 425, 433, 445, 452, 454, 456, 639-640; Moon, 623
Permeability, 980
Permian-Triassic mass extinction, 11, 284, 515
Perovskite, 126, 163, 187, 189
Perry, R. E., 484
Perseid meteor showers, 295, 497
Perturbation, 980
Peru-Chile Trench, 131
Peters, C. A. F., 497
Peterson, Laurence E., 794
Petrography, 980
Phobos (moon), 8, 437-441, 620, 706; craters, 438
Phobos missions, 425, 439, 444
Phoebe (moon), 276, 701-710
Phoebe Regio (Venus), 931
Phoenix. *See* Mars Phoenix
Photochemistry, 980
Photographic emulsions, 74
Photographic imaging, 559-561
Photometers, 980
Photometry, 13, 560, 980
Photomultipliers, 980
Photon energy, 249
Photons, 739, 792, 869, 980
Photopolarimeter, 980
Photosphere, 69, 713-717, 736, 741, 743, 748-751, 756, 980. *See also* Stellar photospheres
Photosynthesis, 110, 357, 565, 980
Photovoltaic cells, 980
Piazzi, Giuseppe, 6, 40, 85, 572
Pierce, A. Keith, 753
Pilkington, John D. H., 653
Pioneer missions; asteroids, 14; Enceladus, 704; Ganymede, 246; Io, 306-307; Jovian passage, 216; Jupiter, 315, 321, 328, 336, 341, 591; Saturn, 591, 621, 674, 681, 685-686, 690, 693, 699; soft particles, 912; solar wind, 789; Venus, 639
Pioneer Venus missions, 45, 239, 631, 917, 933, 940
Pitch, 980
Pixels, 980
Plages (solar), 715, 733
Plains, 985
Planck distribution for thermal radiators, 560
Planck length, 980
Planck time, 26, 29, 868
Planck's constant, 201, 980
Planck's law, 202
Planet (defined), 572-575, 981
Planet-C mission, 921
Planet X theory, 510-516
Planetary atmospheres, 564-571
Planetary classifications and definitions, 572-575, 647
Planetary interiors, 582-588
Planetary migration, 333
Planetary motion, 90; Kepler's laws, 597
Planetary nebulae, 272, 302, 671, 807, 818, 945
Planetary orbits, 594-601; couplings and resonances, 601-605
Planetary ring systems, 605-611
Planetary rotations, 612-617
Planetary satellites, 963, 984
Planetesimals, 15, 88, 108, 115, 141, 280-281, 576, 578, 774, 981; impacts of, 279-286
Planetology; comparative, 630-641; terrestrial planets, 637-641
Planets; classifications and definitions, 85-89, 211; composition, 576; extrasolar, 227-231; formation, 41, 43, 575-582, 667; interiors, 311-312; natural satellites, 617-625; ring systems, 579
Plasma, 981
Plasma sheaths, 981
Plasma sheets, 19, 23, 160, 590
Plasmaspheres; Earth, 591
Plastic deformation, 345; Callisto, 347; Ganymede, 347
Plate boundaries, 130, 626
Plate tectonics, 100, 104, 119, 122, 142, 144, 163, 165, 180, 186, 189, 587, 625-630, 981; Earth, 583; Ganymede, 247; Io, 306; magnetic reversals, 150; Mars, 447; Venus, 584, 916, 935-936
Plato, 599
Pleiades, 301
Plumes (Triton), 858
Plutarch, 495
Plutinos, 86, 353, 619, 634, 648

Pluto (dwarf planet), 352, 510, 513, 572, 597, 622, 634, 641-649, 774; atmosphere, 564, 569; classification controversy, 211; discovery, 647; interior, 585; mass, 598; planetary status, 85-89; rotation, 615
Plutoids, 212, 648
Polar caps; Earth, 95; Ganymede, 245; Mars, 401, 417, 428-437, 452, 455, 632, 635, 639-640, 786; Pluto, 643; Triton, 859
Polar Lander. *See* Mars Polar Lander
Polar orbits, 981
Polarimeters, 981
Polarimetry, 14, 229, 561
Polaris (star), 180
Polarization of light, 561
Poles; defined, 981; magnetic, 146, 154; rotational, 146
Polygonal ground, 981
Polymorphism, 981
Pontecorvo, Bruno, 539
Popper, D. M., 944
Porco, Carolyn, 532
Pores (solar), 745, 812
Portia (moon), 897
Posigrade (defined), 981
Positrons, 981
Potapenko, G. W., 664
Potassium, 768
Potential energy, 981
Pouillet, Claude, 752
Pound, 981
Pound, R. V., 798
Precession, 45, 66, 981; Earth, 180, 613
Price, Buford, 75
Primary minerals, 981
Primary (P) waves. *See* P waves
Primeval atom theory, 871, 875
Primordial soup; Earth, 357; Titan, 856
Prinn, Ronald, 319
Procyon B (star), 942
Prograde orbits, 981
Project Constellation. *See* Constellation program
Project Mohole, 138, 163
Prometheus (moon), 707
Prometheus project, 216
Prominences. *See* Solar prominences
Proper motion, 66-68, 981
Proplyds, 143, 651, 667
Proportional counters, 949
Prospero (moon), 907
Proteins, 231, 981
Proteus (moon), 536
Protogalaxies, 81, 870
Proton precession magnetometers, 152
Proton-proton chain reaction, 723, 726, 848
Protons, 20, 27-28, 73, 106, 156, 982
Protoplanetary disks, 651, 667
Protoplanets, 115, 578-579, 769-770, 982
Protostars, 271, 290, 302, 398, 576, 649-652, 670, 805, 982
Proto-Sun, 108, 115, 368, 372, 576, 847
Protosuns, 982
Proxima Centauri (star), 666
Psamathe (moon), 536
P-T boundary. *See* Permian-Triassic mass extinction
Ptolemy, 68, 599, 616
Pulsar planets, 226
Pulsar timing, 228
Pulsars, 221, 228, 544-545, 578, 653-658, 662, 809, 819, 982; Crab nebula, 257, 819
Purcell, Edward Mills, 303, 664
Pyramids, 3
Pyrolitic Release Experiment (Viking mission), 234
Pyroxene, 116, 769, 982
Quadrantid meteor shower, 366
Quantum mechanical tunneling, 670
Quantum mechanics, 251, 553; and relativity, 877; white dwarf stars, 943
Quantum theory, 258, 982
Quaoar (cubewano), 88, 212, 647
Quarks, 27, 29, 80, 960, 982
Quasar 3C273, 196
Quasars, 28, 30, 291, 546, 548, 662, 982
Queloz, Didier, 222, 229

R-process, 552
Rabinowitz, David, 210
Radar, 982
Radial velocity, 982; detection, 228
Radiant, 982
Radiation; Jupiter, 335-341; nonthermal, 197-200, 660; solar, 751-756; stellar, 398
Radiation belts; Earth, 910-914; Jupiter, 335; Uranus, 894; Venus, 917. *See also* Van Allen radiation belts
Radiative diffusion, 739, 744, 756, 764
Radio astronomy, 236, 659-665, 983; Milky Way, 504; search for extraterrestrial life, 711; solar, 756-759
Radio Astronomy Laboratory, 712
Radio bursts, 757
Radio emissions; Jupiter, 335, 338; Neptune, 529
Radio noise, 983
Radio spectrum, 983
Radio telescopes, 659, 661, 983
Radio waves, 544; auroras, 20; solar disturbances, 592. *See also* Radio astronomy
Radioactive dating; meteorites, 499
Radioactive decay, 983

Radioactive decay clocks, 475, 479
Radioactivity, 983
Radiogenic heating, 983
Radiometric dating, 105
Rain; acid, 511; Mars, 420; primordial, 171, 174; Titan, 235, 707, 852-853; Venus, 640, 916, 926
Rampart craters, 422
Ranger probes, 363, 366, 376
Raup, David, 284, 514
Ray, Ernest C., 912
Ray craters, 347
Real time, 983
Reber, Grote, 664
Rebka, G. A., 798
Reconnaissance Orbiter. *See* Mars Reconnaissance Orbiter
Reconnaissance satellite, 983
Red dwarf stars, 31, 220, 222, 513, 665-669, 983
Red giant stars, 271, 302, 399, 544, 669-673, 723, 807-808, 945, 983; Sun as, 98, 745
Red Planet, 983. *See also* Mars
Red Spot. *See* Great Red Spot; Little Red Spot
Red supergiant stars, 272, 669-673, 724, 807; supernovae, 820. *See also* Supergiant stars
Redshifting, 24, 28, 79, 561-562, 871, 983; cosmological, 876-877, 881; gravitational, 257. *See also* Doppler shifts
Reentry, 983
Reese, Elmer, 318, 326
Reflecting telescopes, 558, 561, 983
Reflection, 983
Reflection nebulae, 301
Reflectivity, 439, 983
Refracting telescopes, 983
Refraction, 983
Refractor telescopes, 558
Refractory elements, 984
Refractory substances, 619
Regolith, 984; asteroids, 10; lunar, 366, 370-371, 384-389, 396
Reines, Frederick, 538
Relativity, 984; general, 969; Christiaan Huygens, 254
Relativity theory, 24, 78, 600, 732, 760, 851, 871, 875, 880; degenerate electrons, 944; general, 82, 196, 248-252, 257; and quantum mechanics, 877; space-time, 795-800, 801-805; special, 248; testing, 944
Remanent magnetism, 145-147, 152-153
Remote sensing, 984
Resolution, 984
Resonances, 7, 601-605
Respiration, 110
Retrograde orbits, 984
Retrograde rotation, 914
Reuven Ramaty High Energy Solar Spectroscopic Imager, 951
Revelstoke (meteorite), 473
Revolution (defined), 984
Reynolds, Ray, 587
Rhea (moon), 610, 703, 705, 708
Rhea Mons (Venus), 931
RHESSI. *See* Reuven Ramaty High Energy Solar Spectroscopic Imager
Ribonucleic acid, 231, 356; origins, 357
Riccioli, Giambattista, 360
Richer, Jean, 181
Rickman, Hans, 45
Riemann, Georg Friedrich, 796, 798
Ries crater, 280
Rigel (star), 822
Right ascension, 984
Rigid body motion, 589
Rilles, 984
Ring resonance, 608, 984
Ring systems, 343, 579, 605-611, 635; Jovian moons, 314; Jupiter, 341-344; Neptune, 530-535, 537; resonances, 601; Saturn, 603, 688, 693-700; Uranus, 896-901
Rivers; lava, 627; Mars, 417, 445, 458
RNA. *See* Ribonucleic acid
Robotics, 984
Roche, Édouard, 605, 699
Roche limit, 341, 437, 605, 984
Rock cycle, 129
Rodgers, John, 437
Roentgen satellite, 54
Roll, 980
Roll, P. G., 26
Roman, Nancy, 833
Rømer, Ole, 195
Rosalind (moon), 897
ROSAT. *See* Roentgen satellite
Rosetta mission, 51, 61, 63
Rotation, 612-617, 984; Earth, 177-183, 612; Europa, 216; Jovian planets, 312; Jupiter, 183, 318, 332, 614; Mars, 614; Mercury, 183, 613; Neptune, 615; Pluto, 615; Saturn, 183, 615; solar, 734, 741, 744-745, 753, 761; Uranus, 615; Venus, 183, 613, 914
Rovers. *See* Lunokhod rovers; Mars Exploration Rovers; Sojourner
RR Lyrae stars, 544
Rubbia, Carlo, 29
Russell, Henry Norris, 269, 273, 553, 726
Rutherford, Ernest, 726
Ryle, Martin, 661

S-process, 551
S-type asteroids, 9
S-wave shadow zone, 144
S waves, 123, 126-127, 143, 163-164, 188, 984-985
Sagan, Carl, 236, 339, 404, 926

Sagittarius A, 504
Sagittarius B2, 301
Sakigake probe, 50
Salpeter, Edwin, 553
Sampson, R. A., 216
San Andreas fault, 131
Sandage, Allan, 661
Sao (moon), 536
Sapping, 442-443, 455
SAR arcs, 20-21
Saros cycle, 194, 984
Satellites (defined), 984
Satellites (natural), 617-625; formation, 774; icy, 585; Jovian planets, 314; tidal forces, 585. *See also specific names*
Saturn, 33, 115, 235, 311-317, 621, 633; atmosphere, 112, 564, 568, 615, 674-682; auroras, 21; interior, 586, 683-686; relationship with Jupiter, 334; magnetic field, 148, 686-692; magnetosphere, 591; moons, 701-710; oblateness, 312, 674, 683; radio emissions, 688-689; rings, 603, 605-611, 633, 693-700; rotation, 615; satellites, 621, 624, 634. *See also* Jovian planets
Sausage ring, 537
Scalar-tensor theory, 732
Scarps, 984
Scattered disk objects, 86, 353; Eris, 208-213
Schiaparelli, Giovanni, 401, 428, 495, 497, 604
Schröter, Johann, 465
Schwabe, Samuel Heinrich, 743, 811
Schwarzschild, Karl, 249, 272, 798-799
Schwarzschild, Martin, 809
Schwarzschild radius, 798-799, 984
Scintillation, 984
Scintillation counters, 75, 538, 949, 984
Scorpius X-1, 546-547
Scott, Paul Frederick, 653
ScP waves, 127
ScS waves, 127
SDOs. *See* Scattered disk objects
Seafloor spreading, 132, 150
Search for extraterrestrial intelligence, 710-713
Search for Extraterrestrial Intelligence (SETI), 236, 399, 662
Seasons; Earth, 95-97, 177, 612, 616; Mars, 614; Saturn, 615, 676; Triton, 859; Uranus, 615, 830
Seawater, 166-172
Secchi, Pietro Angel, 401
Secondary craters, 984
Secondary (S) waves. *See* S waves
Secular variations, 126, 149-155
Sedimentary rocks, 129-130, 984; Venus, 927, 932
Sedna (scattered disk object), 88, 210, 212, 556, 647
See, Thomas, 227
Seismic activity, 985
Seismic refraction studies, 136
Seismic tomography, 124, 128, 165, 189
Seismic waves, 119, 121, 123-124, 127-128, 132-133, 147, 163-164, 187, 189, 985; lunar, 375
Seismology, 123, 127; lunar, 395, 587; planetary interiors, 587
Seismometers, 985
Selene mission, 364
Selenography, 985
Selenology, 985
Self-exciting dynamos, 985
Sepkoski, J. John, Jr., 284, 514
SETI. *See* Search for Extraterrestrial Intelligence
SETI Institute, 711-712
70 Virginis planet, 222-223
Seyfert galaxies, 985
Shane telescope, 825
Shapley, Harlow, 303, 501, 726, 883-884
Shatter cones, 282
Shear (S) waves. *See* S waves
Shenzhou missions, 388
Shepherd satellites, 314, 532, 534, 537, 579, 607, 609-610, 633, 696, 707, 898-899, 903-904, 985
Shepherding moons. *See* Shepherd satellites
Shergottites, 470
Shield volcanoes (Venus), 938
Shock metamorphism, 282
Shock waves, 985
Shoemaker, Carolyn, 52
Shoemaker, Eugene, 52, 363
Shoemaker-Levy 9 (comet), 52-56
Shu, Frank, 724
SI units, 985
Siberian Traps, 12, 284, 515
Siberian Tunguska event, 64, 497
Side-slip plate boundaries, 130
Sidereal day, 178
Sidereal period, 612, 985
Sidereal month, 90
Siderites, 481
Siderophile elements, 984
Siderophyres, 488
Sif Mons (Venus), 939
Silicates, 186, 582, 587, 985
Singularities, 985
Sinuous rilles, 379, 441, 985
Sirius B (white dwarf), 942
SIRTF. *See* Spitzer Space Telescope
Skylab missions, 719, 779, 781, 863, 953; solar data, 71, 745-746
SL 9. *See* Shoemaker-Levy 9.
Slipher, Earl C., 402
Slipher, Vesto Melvin, 874

Small Astronomy Satellites, 951
Small bodies; composition, 771; interiors, 584. *See also* Asteroids; Comets; Cubewanos; Dwarf planets; Planets; Plutinos; Scattered disk objects; Small solar-system bodies
Small Missions for Advanced Research in Technology, 364, 623
Small solar-system bodies, 87. *See also* Asteroids; Comets; Cubewanos; Dwarf planets; Planets; Plutinos; Scattered disk objects
SMART-1. *See* Small Missions for Advanced Research in Technology
SMM. *See* Solar Maximum Mission
Smog, 852; Titan, 856; Triton, 858
Smokers (hydrothermal vents), 168, 985
Smooth plains, 985
SNC meteorites, 470, 492
Snell's law, 133, 985
Sodium, 768
SOHO. *See* Solar and Heliospheric Observatory
Soil, 358; lunar, 386-387
Sojourner (Mars Rover), 410
Solar and Heliospheric Observatory, 71, 298, 716, 721, 725, 762
Solar array, 986
Solar cell, 986
Solar constant, 260, 751-756, 784, 986
Solar cycle, 297, 720, 743, 749, 753, 811-815, 986; solar constant and, 753
Solar day, 178, 612
Solar Dynamics Observatory, 758
Solar eclipses, 192, 196, 250
Solar energy, 850
Solar flares, 17, 97, 156, 160-161, 297, 721, 727-732, 745, 757, 780, 792-795, 912, 986; classification, 730; damage from, 298; radio bursts, 757; surges, 728
Solar magnetic field, 159
Solar mass, 986
Solar maximum, 297
Solar Maximum Mission, 720, 729-730, 736, 746, 753-754, 781, 785, 793, 863, 951, 953
Solar nebula, 104, 115, 140, 218, 372, 575, 773; rings and, 343
Solar neutrino problem, 542
Solar neutrino units, 540
Solar Orbiter, 758
Solar oscillations, 732-736, 759-763
Solar prominences, 97, 715, 727, 745, 780-782, 981
Solar radiation, 986
Solar Radiation I, 730
Solar storms. *See* Coronal mass ejections
Solar system, 986; composition, 582, 773; element distribution, 767-773; evolution, 610; origins, 104, 108, 115, 768, 773-778
Solar wind, 19, 71, 108, 156, 159, 167, 296, 589, 718, 721, 787-792, 986; auroras, 19; dangers of, 298, 592; formation, 776; effect on Jupiter, 337; magnetic field, 745; planet formation, 577; effect on Saturn, 688
Solstices, 96
Sonar, 986
Sounding sensor, 986
South magnetic pole, 155
South Tropical Zone (Jupiter), 326
Space age, 986
Space Infrared Telescope Facility. *See* Spitzer Space Telescope
Space Interferometry Mission, 669
Space telescopes, 832-843, 986
Space-time, 248-252, 257, 986; curvature of, 249; distortion by gravity, 795-800; mathematics, 801-805
Space Vision for Exploration, 372, 382, 393
Space weather, 72, 298
Spacecraft, 986
Spacelab, 22
Spallation, 75
Special relativity, 248, 801, 803, 984, 989
Specific gravity, 312, 986
Specific heat, 986
Spectra, 986
Spectral class, 986
Spectral shifts. *See* Blueshifting; Doppler shifts; Redshifting
Spectrographs, 986
Spectrometers, 986
Spectrometry, 169; infrared, 972
Spectroscopes, 986
Spectroscopy, 14, 559-561, 645, 771, 862, 986; solar, 714
Spectrum-luminosity diagram, 269
Speed of light, 801
Spherules, 987
Spicules, 715, 764, 780
Spin, 987
Spin axis, 987
Spin-flip transitions, 197-198, 502-503, 660
Spin-orbit coupling, 267
Spin-orbit resonance, 601
Spinel, 163, 187
Spinrad, Hyron, 824
Spiral arms, 987
Spiral galaxies, 501-503, 874, 884, 987
Spirit (Mars Rover), 235, 410, 412. *See also* Mars Exploration Rovers

Spitzer, Lyman, Jr., 289, 833, 866
Spitzer Space Telescope, 62, 224-225, 289, 580, 834, 837, 841
Spörer minimum, 745, 785
Spots. *See* Dark Spots; Great Red Spot; Little Red Spot
Sputnik 1, 809, 912
Sputnik 3, 912
Sputtering, 987
Square Kilometer Array, 829
Stagnation point, 790, 987
Standard candles. *See* Candles
Standard Solar Model, 761
Star trackers, 987
Stardust mission, 51, 60-61, 233
Starfish project, 912
Stark, Johannes, 944
Stark effect, 944
Stars, 987; evolution, 269-274, 543-549, 657; formation, 81; nuclear synthesis, 549-554; red dwarfs, 665-669
Starspots, 814
Steady state theory, 29-30, 79, 875, 987
Steady state universe, 875
Stebbins, Robin, 733
Stefan, Jožef, 201
Stefan-Boltzmann law, 201-202, 269, 987
Steinback meteorite, 488
Stellar cores, 987
Stellar evolution, 805-811, 987
Stellar luminosity, 987
Stellar-mass black holes, 250
Stellar photospheres, 987
Stellar populations, 987
Stellar wind, 303, 670, 987
Stelliferous Age, 666
Stokes, George Gabriel, 513
Stonehenge, 1
Stony-iron meteorites, 9, 115, 143, 468, 486-490
Stony meteorites, 115, 468, 776
Stratosphere (Earth's), 565, 987
Streamers, 720
String theory, 258, 882
Strobel, Darrel, 319
Strömgren, Bengt, 273
Strong nuclear force, 987
Struve, Friedrich Georg Wilhelm von, 48, 882
Stukeley, William, 1
Sub-auroral red arcs, 20-21
Subduction, 131, 142, 165, 187
Subduction zones, 626
Sublimation, 987
Subsatellites, 987
Substellar objects. *See* Brown dwarf stars
Sudbury Neutrino Observatory, 400
Suisei probe, 45, 50
Sulfur; Venus, 927
Sulfur dioxide; Earth, 109; Venus, 915, 940
Summer solstice. *See* Solstices
Sun; active, 749; chromosphere, 69, 713-717, 736, 757, 764; core, 739-742, 764; corona, 69-73, 589, 714, 717-723, 737, 765, 779-780, 794; coronal loops, 70; effect on Earth, 94-99, 721; evolution, 723-727; formation, 104, 774; geodesy, 732-736; gravitational field, 296; gravitational waves, 734; infrared emissions, 286, 736-739; interior, 739-742; magnetic fields, 70, 148, 727, 743-748; nucleosynthesis in, 549-554; photosphere, 69, 713, 736, 743, 748-751, 756; prominences, 97, 715, 727, 745, 780-782; quiet, 749; radiation, 751-756; radio emissions, 756-759; red giant fate, 670; rotation, 745, 761, 789; seismology, 759-763; solar eclipse, 192; solar flares, 727-732; spectrum, 202; structure, 763-767; temperature, 202, 737; thermonuclear reactions, 847-852; ultraviolet emissions, 778-784, 986; ultraviolet images, 864; variability, 784-787; X-ray emissions, 792-795
Sundials, 181
Sunspot cycle, 71, 715, 728, 730, 743, 749, 753, 784, 811-815
Sunspot pairs, 744
Sunspots, 17, 97, 156, 720, 727, 730, 733, 743-744, 753, 780-782, 784, 792, 811-815, 988
Super Earths, 224
Superclusters. *See* Clusters and superclusters
Superconducting magnetometers, 152
Supercritical fluid, 331
Superfluids, 988
Supergiant stars, 270, 544. *See also* Red supergiant stars
Supergiants, 988
Supergranules, 749, 765
Superimposition, 988
Superior conjunction, 962
Super-massive black holes, 251
Supernova 1987A, 541-542, 656, 820, 953
Supernova remnants, 988
Supernovae, 81, 242, 250, 272, 303, 539, 544-545, 548, 551, 656, 807, 816-823, 849, 988; neutrinos, 541-542; origin of the solar system, 373; planet formation, 575; pulsars and, 655-656; as standard

candles, 877; Type II, 672; ultraviolet research, 865; white dwarfs, 945
Surt (Io), 307
Surveyor missions, 363, 376, 384, 394
Swedenborg, Emanuel, 724
Swift mission, 238, 240, 242, 837, 841, 951
Swift-Tuttle (comet), 497
Swing-bys, 988
Synchronous relationships, 213, 244, 275, 858; Uranian satellites, 902
Synchronous rotation, 602, 620, 988
Synchrotron radiation, 198, 655-656, 757; Jupiter, 335
Synodic months, 90, 193, 975
Synodic periods, 988
Syrtis Major, 401
Système International d'Unités. *See* SI units

T-Tauri stage, 723, 770, 988
T-Tauri stars, 650
Tachocline, 740
Taenite, 481, 488, 988
Tails (cometary), 988
Tarter, Jill, 31
Tau Bootis (star), 222
Taylor, Donald J., 653
TD-1 (satellite), 864
Tectonics, 625-630; Venus, 923, 939. *See also* Plate tectonics
Teide (brown dwarf), 32
Tektites, 512, 988
Telemetry, 988
Teleoperations, 988
Telescopes, 832, 988; Cassegrain, 961; ground-based, 824-831; Kirkpatrick-Baez, 948; optical, 558; radio, 983; reflecting, 983; refracting, 983; space-based, 832-843, 986; Wolter, 947
Telesto (moon), 603
Television waves, 544
Teller, Edward, 25
Tempel 1 (comet), 51, 61-62
Tempel 9P (comet), 353, 557
Tempel-Tuttle (comet), 497
Temperatures; Earth, 95, 260; Jovian planets, 311; Moon, 95; outer planets, 95; Venus, 95
Termination shock boundary, 296
Terminators, 988
Terracing (Mars), 448
Terrestrial planets, 218, 311, 524, 774, 843-847, 988. *See also* Earth; Mars; Mercury; Venus
Tesserae (Venus), 930
Tethys (moon), 603, 703, 708
Thalassa (moon), 536
Tharsis Dome (Mars), 447-448, 988
Tharsis Ridge (Mars), 585, 628
Thea Mons (Venus), 931
Thebe (moon), 341
Theories of everything, 26, 29, 82, 251, 258, 877. *See also* Grand unified theories; Unified field theories
Thermal emissions, 197, 201-204, 660
Thermal erosion, 988
Thermal mapping, 988
Thermokarst, 988; Mars, 455-456
Thermonuclear reactions, 847-852, 988. *See also* Fusion reactions; Nucleosynthesis
Thermosphere (Earth's), 989
Theta auroras, 19
Thirty Meter Telescope, 828
Tholins (Pluto), 210
Thom, Alexander, 5
Thomson, Sir William. *See* Kelvin, Lord
Thorne, Kip, 548
Tidal flexing, 245, 858
Tidal forces, 91, 585, 616, 989; Jovian, 306; Jupiter, 628
Tidal friction, 98
Tidal locking; Earth's moon, 602; Jovian satellites, 350, 602; Martian satellites, 602; red dwarf systems, 668
Tides; lunar, 178; ocean, 183
Tiger stripes (Enceladus), 205
Time dilation, 257, 989
Time measurement, 181
Timocharis, 465
Tired light hypothesis, 82
Titan (moon), 314-316, 621, 634, 701-710, 852-857; atmosphere, 569, 592; density, 579; life possibilities, 235; magnetic field, 689
Titania (moon), 903
Titanium, 768
Titius, Johann, 6, 40, 572
TNOs. *See* Trans-Neptunian objects
TOE. *See* Theories of everything
Tomasko, Martin, 319
Tombaugh, Clyde, 85, 572, 641, 647
Tomography. *See* Seismic tomography
Tonk (meteorite), 473
Topographic maps, 153
Topside observation, 989
Torino scale, 12
Toro (asteroid), 14
Torsion balance, 255
Total lunar eclipse, 191
Total solar eclipse, 192, 194, 250
Tousey, Richard, 719
TRACE. *See* Transition Region and Coronal Explorer
Transfer orbits, 989
Transform plate boundaries, 130
Transit detection method, 228
Transition region (Sun), 779, 781-782
Transition Region and Coronal Explorer, 72

Transits, 193, 989
Trans-Neptunian objects, 86
Transponders, 989
Tree rings, 745
Triangulation, 989
Trifid nebula, 300
Triple-alpha process, 550
Triton (moon), 316, 535-536, 615, 621-622, 634, 647, 857-861; life possibilities, 235; retrograde orbit, 858; volcanism, 629
Troilite, 768, 989
Trojan asteroids, 8-10, 602-603
Tropopause (Earth's), 565
Tropospheres, 989; Earth, 565
Trujillo, Chad A., 210
Trumpler, Robert Julius, 303
Tunguska event, 64, 497
Tunneling, 989
Twin paradox, 803
Twining, A. C., 497
2-Micron All Sky Survey, 666
Tyrrhena Patera (Mars), 449

U-type asteroids, 9, 14
Udry, Stéphane, 224
Uhuru. *See* Cosmic X-Ray Explorer
UKIRT. *See* United Kingdom Infrared Telescope
Ultramafic rocks, 975
Ultraviolet astronomy, 718, 862-867, 989
Ultraviolet radiation, 544, 862, 989; solar, 778-784, 986
Ultraviolet spectrometers, 989
Ulysses mission, 758
Umbra, 989; lunar, 192; sunspots, 812
Umbral lunar eclipse, 191
Umbriel (moon), 902
Unified field theories, 26, 251. *See also* Grand unified theories; Theories of everything
Uniformitarianism, 104, 990
United Kingdom Infrared Telescope, 288
Universal gravitation law, 253; equation, 253
Universe; accelerating expansion, 876; age of, 30; end state, 548; evolution, 867-874; expansion, 874-879; oscillating, 30; structure, 879-886
Upsilon Andromedae (star), 222
Uranus, 115, 311-317, 634; atmosphere, 112, 564, 568, 634, 886-890; auroras, 21; clouds, 292; formation, 578; interior, 586, 890-893; magnetic field, 893-896; magnetosphere, 592; rings, 196, 605-611, 896-901; rotation, 615; satellites, 621, 624, 901-905; seasons, 830; tilt, 906-909. *See also* Jovian planets
Ureilites, 469
Urey, Harold, 232
Urey-Miller experiment. *See* Miller-Urey experiment
Ussher, James, 104
Uxmal, 3

Vacuum, 990
Valles Marineris (Mars), 402, 585, 628
Van, James A., 161
Van Allen, James A., 156, 158, 338, 794, 912
Van Allen radiation belts, 156, 158, 161, 316, 338, 547, 591-592, 729, 790, 910-914, 990
Van de Kamp, Peter, 227
Van der Meer, Simon, 29
Van Maanen, Adriaan, 666, 942
Van Maanen's star (white dwarf), 942
Variable stars, 544, 666, 990; Sun, 784-787
Varuna (cubewano), 88
Vastitas Borealis (Mars), 457
Vega missions, 44, 50, 631, 917, 927
Vega system, 290
Vela missions, 238, 951
Vela pulsar, 653, 662
Velikovsky, Immanuel, 402
Venera missions, 239, 630, 639, 921, 925, 927, 933; Venus, 917
Venus, 112, 115, 630, 634, 844; albedo, 261, 566; atmosphere, 108, 111, 415, 564-565, 631, 914-921; craters, 921-925; greenhouse effect, 232; interior, 584; ionosphere, 590; magnetic field, 148; magnetosphere, 590; Mayan significance, 3; rotation, 613; soil and rock samples, 925-929, 933; surface, 639, 929-938; tectonics, 627; temperatures, 95; as terrestrial planet, 637, 843-847; volcanoes, 938-941
Venus Express, 918, 921, 928, 935
Verhoogen, John, 169
Vernal equinox, 66-67, 990
Vernazza, Jorge, 753
Vernier rocket, 990
Very Large Array, 829, 840
Very Large Telescope, 826
Very long baseline interferometry, 662, 990
Vesta (asteroid), 7, 14, 41, 88, 469; Dawn mission, 15; discovery, 40
Via galacta, 303
Vico, Francisco di, 48
Viking missions, 409, 449; Cydonian plain, 403; landers, 450; life experiments, 233-234, 424, 635; Martian seasons, 454; Martian soil, 407, 416, 431, 444, 456; Martian temperatures, 412, 429,

442, 452; Phobos and Deimos, 437, 439; surface images, 640
Viking satellite (Sweden), 22
Virgo supercluster, 879
Virial theorem, 990
Viscosity, 990
Visible light, 544, 558, 659, 717, 751, 832; colors, 752
Visible rays, 990
Vision for Space Exploration, 387, 623
Visual binary stars, 990
VLBI. *See* Very long baseline interferometry
Void, 990
Volatiles, 58, 108, 167, 619, 990
Volcanic vents, 357
Volcanism, 100, 102, 134, 165, 703, 990; atmosphere formation, 415; Earth, 117; Hawaiian islands, 164; ice, 585; Io, 219, 304-310, 345, 586, 591, 628; lunar, 362, 366, 378, 390-391, 396; Mars, 423, 443, 446-452, 628, 631, 639; mountain building, 132; outgassing, 168; terrestrial planets, 635; Titan, 854-855; Triton, 858; Venus, 627, 639, 915, 922-923, 931-932, 934, 938-941; water, 347, 453
Voyager missions; auroras, 21, 23; Cassini division, 604; Enceladus, 205; Europa, 214, 217; Ganymede, 246; Great Red Spot, 328; Iapetus, 277; interplanetary environment, 296; Io, 305, 307, 309, 629; Jovian planets, 311, 314, 316; Jupiter, 321, 341, 345, 591; Jupiter's rings, 606, 633; Jupiter's satellites, 621; Mercury, 466; Miranda, 579; Neptune, 313, 518, 528, 535, 569, 634; Neptune's rings, 609; Saturn, 674, 678-679, 685, 690; Saturn's rings, 693, 697, 699; Saturn's satellites, 707-708; Titan, 701; Uranus, 634, 887, 889, 900
Vulcan theory, 760, 798

Walker, Robert, 75
Wallace, Alfred Russel, 402
Water; Charon, 644; cometary, 46; Earth, 95, 166-172; Enceladus, 205; Europa, 216, 218-219, 346, 350; Ganymede, 244, 246, 346, 350; HD 189733b, 224; ice, 587; Jovian satellites, 585; life, 266, 667; lunar, 364, 397, 619; Mars, 406, 409-410, 412-413, 442, 452-459, 632, 639-640; meteorites, 578; Nereid, 536; outer planets, 585; Phobos and Deimos, 438-439; protoplanetery disks, 293; solar-system formation, 774; star systems, 292; terrestrial planets, 415; vapor, 109, 261; Venus, 934, 936
Water of hydration, 990
Waterson, John James, 725
Watson-Watt, Sir Robert A., 663
Wavelength, 990
Weakly interacting massive particles, 882
Weather; Earth, 565; Jovian planets, 311, 316; solar wind and, 593. *See also* Space weather
Weathering; processes, 130; Venus, 926
Webb Space Telescope. *See* James Webb Space Telescope
Weber, Joseph, 255
Wegener, Alfred, 587, 625
Weird Terrain (Mercury), 462, 602
Weizsäcker, Carl von, 581, 724, 851
West, Robert, 320
Westerbork Array, 829
Wheeler, John, 250, 547, 797, 819
Whipple, Fred L., 44, 57, 498
White dwarf stars, 257, 271-272, 551, 655-656, 671, 724, 807-808, 818, 849, 942-946, 990; novae, 816; supernovae, 818
White-light flares, 728
White spots, 316, 990
Whitmire, Daniel, 515
Widmanstätten, Alois Josep, 481
Widmanstätten patterns, 481, 990
Wien, Wilhelm, 201
Wien's law, 201-202
Wild 2 (comet), 51, 60, 62
Wilkinson, David T., 26
Wilkinson Microwave Anisotropy Probe, 26, 82, 881
Willson, Richard C., 754
Wilson, Robert W., 26, 82, 871, 875
Wilt, Rupert, 737
WIMPS. *See* Weakly interacting massive particles
Winckler, John, 794
Wind mission, 864
Wind tunnels, 990
Winds; Earth, 178; Mars, 567, 632; Triton, 858; Uranus, 569, 888; Venus, 915, 918, 932
Winter solstice. *See* Solstices
Wisdom, Jack, 222
WMAP. *See* Wilkinson Microwave Anisotropy Probe
Wolf, Max, 7-8, 49
Wolf 489 (white dwarf), 944
Wolfe, John H., 337
Wolszczan, Alexander, 221, 228

Wolter telescopes, 837, 947, 951
Wood, John, 1
Work, 991
World line, 795, 797, 801, 991
Worlds in Collision (Velikovsky), 402

X-band, 991
X-radiation, 991
X-ray astronomy, 547-548, 947-954, 991; solar research, 71, 725
X-ray bursters, 546, 793
X-ray sources, 543-549
X-ray stars, 243
X rays, 18, 189, 544-545, 548, 862, 947; solar, 792-795. *See also* X-ray astronomy
Xena. *See* Eris
Xenoliths, 136
Xenon, 789
XMM-Newton observatory, 841

Yaw, 980
Year (defined), 96
Yellow giant stars, 668
Ylem, 25
Yohkoh mission, 71

Zach, Franz von, 40
Zeeman, Pieter, 743
Zeeman effect, 561, 811, 991
Zero gravity, 991
Zircon crystals, 117, 119, 142
Zircons, 991
Zodiacal light, 295, 991
Zones, 991
Zupus, Giovanni Battista, 465
Zwicky, Fritz, 547, 662