Plasma Physics of the Local Cosmos

Committee on Solar and Space Physics
Space Studies Board
Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by Contracts NASW-96013 and NASW-01001 between the National Academy of Sciences and the National Aeronautics and Space Administration.

International Standard Book Number 0-309-09215-9 (POD)
International Standard Book Number 0-309-53215-9 (PDF)

Cover—Top: The aurora australis (southern lights) photographed from the International Space Station on April 18, 2003. Courtesy of Earth Sciences and Image Analysis Laboratory, NASA Johnson Space Center. Bottom: Conceptual representation of the heliosphere and the solar system's immediate galactic environment. Distances in astronomical units (AU) are indicated on a logarithmic scale. (1 AU is the mean distance between the Sun and the Earth, or roughly 150,000,000 kilometers.) Courtesy of P. Liewer (Jet Propulsion Laboratory) and R. Mewaldt (California Institute of Technology).

Copies of this report are available free of charge from:

Space Studies Board
National Research Council
500 Fifth Street, N.W.
Washington, DC 20001

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Wm. A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. Wm. A. Wulf are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

OTHER REPORTS OF THE SPACE STUDIES BOARD

Issues and Opportunities Regarding the U.S. Space Program: A Summary Report of a Workshop on National Space Policy (2004)

"Assessment of NASA's Draft 2003 Earth Science Enterprise Strategy" (2003)
"Assessment of NASA's Draft 2003 Space Science Enterprise Strategy" (2003)
Satellite Observations of the Earth's Environment: Accelerating the Transition of Research to Operations (2003)
Steps to Facilitate Principal-Investigator-Led Earth Science Missions (2003)
The Sun to the Earth—and Beyond: Panel Reports (2003)

Assessment of Directions in Microgravity and Physical Sciences Research at NASA (2002)
Assessment of the Usefulness and Availability of NASA's Earth and Space Science Mission Data (2002)
Factors Affecting the Utilization of the International Space Station for Research in the Biological and Physical Sciences (2002)
Life in the Universe: An Assessment of U.S. and International Programs in Astrobiology (2002)
New Frontiers in the Solar System: An Integrated Exploration Strategy (2002)
Review of NASA's Earth Science Enterprise Applications Program Plan (2002)
"Review of the Redesigned Space Interferometry Mission (SIM)" (2002)
Safe on Mars: Precursor Measurements Necessary to Support Human Operations on the Martian Surface (2002)
The Sun to the Earth—and Beyond: A Decadal Research Strategy in Solar and Space Physics (2002)
Toward New Partnerships in Remote Sensing: Government, the Private Sector, and Earth Science Research (2002)
Using Remote Sensing in State and Local Government: Information for Management and Decision Making (2002)

Assessment of Mars Science and Mission Priorities (2001)
The Mission of Microgravity and Physical Sciences Research at NASA (2001)
The Quarantine and Certification of Martian Samples (2001)
Readiness Issues Related to Research in the Biological and Physical Sciences on the International Space Station (2001)
"Scientific Assessment of the Descoped Mission Concept for the Next Generation Space Telescope (NGST)" (2001)
Signs of Life: A Report Based on the April 2000 Workshop on Life Detection Techniques (2001)
Transforming Remote Sensing Data into Information and Applications (2001)
U.S. Astronomy and Astrophysics: Managing an Integrated Program (2001)

Limited copies of these reports are available free of charge from:
Space Studies Board
The National Academies
500 Fifth Street, NW, Washington, DC 20001
(202) 334-3477
ssb@nas.edu
www.nationalacademies.org/ssb/ssb.html

NOTE: Listed according to year of approval for release.

COMMITTEE ON SOLAR AND SPACE PHYSICS

JAMES L. BURCH, Southwest Research Institute, *Chair*
CLAUDIA J. ALEXANDER, Jet Propulsion Laboratory
VASSILIS ANGELOPOULOS, University of California, Berkeley
ANTHONY CHAN, Rice University
ANDREW F. CHENG, Johns Hopkins University
JAMES F. DRAKE, JR., University of Maryland, College Park
JOHN C. FOSTER, Massachusetts Institute of Technology
STEPHEN A. FUSELIER, Lockheed Martin Advanced Technology Center
SARAH GIBSON, National Center for Atmospheric Research
CRAIG KLETZING, University of Iowa
GANG LU, National Center for Atmospheric Research
BARRY H. MAUK, Johns Hopkins University
FRANK B. McDONALD, University of Maryland, College Park
EUGENE N. PARKER, University of Chicago, Professor Emeritus
ROBERT W. SCHUNK, Utah State University
GARY P. ZANK, University of California, Riverside

Staff

ARTHUR CHARO, Study Director
WILLIAM S. LEWIS,[1] Consultant
THERESA M. FISHER, Senior Program Assistant

---

[1] On temporary assignment from Southwest Research Institute.

http://www.nap.edu/openbook/0309092159/html/R6.html, copyright 2004, 2001 The National Academy of Sciences, all rights reserved

SPACE STUDIES BOARD

LENNARD A. FISK, University of Michigan, *Chair*
GEORGE A. PAULIKAS, The Aerospace Corporation (retired), *Vice Chair*
J. ROGER P. ANGEL, University of Arizona
ANA P. BARROS, Harvard University
RETA F. BEEBE, New Mexico State University
ROGER D. BLANDFORD, Stanford University
JAMES L. BURCH, Southwest Research Institute
RADFORD BYERLY, JR., University of Colorado
HOWARD M. EINSPAHR, Bristol-Myers Squibb Pharmaceutical Research Institute (retired)
STEVEN H. FLAJSER, Loral Space and Communications, Ltd.
MICHAEL H. FREILICH, Oregon State University
DON P. GIDDENS, Georgia Institute of Technology/Emory University
DONALD INGBER, Harvard Medical School
RALPH H. JACOBSON, The Charles Stark Draper Laboratory (retired)
TAMARA E. JERNIGAN, Lawrence Livermore National Laboratory
MARGARET G. KIVELSON, University of California, Los Angeles
CALVIN W. LOWE, Bowie State University
BRUCE D. MARCUS, TRW, Inc. (retired)
HARRY Y. McSWEEN, JR., University of Tennessee
DENNIS W. READEY, Colorado School of Mines
ANNA-LOUISE REYSENBACH, Portland State University
ROALD S. SAGDEEV, University of Maryland
CAROLUS J. SCHRIJVER, Lockheed Martin Solar and Astrophysics Laboratory
ROBERT J. SERAFIN, National Center for Atmospheric Research
MITCHELL SOGIN, Marine Biological Laboratory
C. MEGAN URRY, Yale University
J. CRAIG WHEELER, University of Texas, Austin

JOSEPH K. ALEXANDER, Director

Preface

This report originated in 1999 as a result of discussions between the Committee on Solar and Space Physics (CSSP) and officials within NASA's Office of Space Science Sun-Earth Connections program. As noted in the statement of task (Appendix A), the objective of the study was to provide a scientific assessment and strategy for the study of magnetized plasmas in the solar system. By emphasizing the connections between locally occurring (solar system) structures and processes and their astrophysical counterparts, the study would contribute to a unified view of cosmic plasma behavior. An additional objective was to relate basic scientific studies of plasmas to studies of the Sun's influence on Earth's space environment.

The study was under way when the Space Studies Board was asked in early 2000 to conduct a decadal survey in solar and space physics. The CSSP stood down during the next 18 months as all of its members served on either the study's Survey Committee or one of its five study panels. A pre-print of the Survey Committee's report was delivered to agency sponsors in August 2002. The Survey Committee's report and a separate volume containing the reports of the survey's five panels were published in 2003.

While part of the original intent of this study was accomplished by the decadal survey—the Survey Committee and panel reports provide priorities and strategies for future program activities—members of CSSP completed this report to address the other objectives. The present report differs substantially from an initial draft that was completed prior to the commencement of the survey activities. In particular, CSSP defers to the Survey Committee's report for recommendations and endorses those. *The committee views this report as a primer that will provide a unified view of the field and show its connections to other scientific disciplines, especially astrophysics.* The audience for the report includes scientists working in fields outside but related to space physics, graduate students in space physics, agency officials, and interested congressional staff and members of the public.

Acknowledgment of Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

Amitava Bhattacharjee, University of Iowa,
Joachim Birn, Los Alamos National Laboratory,
Timothy E. Eastman, Plasmas International,
J.R. Jokipii, University of Arizona,
Andrew F. Nagy, University of Michigan,
Robert Rosner, University of Chicago, and
Michelle F. Thomsen, Los Alamos National Laboratory.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of this report was overseen by Mihaly Horanyi, University of Colorado. Appointed by the National Research Council, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Contents

APPENDIXES

Executive Summary

Earth's neighborhood in space—the local cosmos—provides a uniquely accessible laboratory in which to study the behavior of space plasmas (ionized gases) in a wide range of environments. By taking advantage of our ability to closely scrutinize and directly sample the plasma environments of the Sun, Earth, the planets, and other solar system bodies, we can test our understanding of plasmas and extend this knowledge to the stars and galaxies that we can view only from afar.

Solar and space physics research explores a diverse range of plasma physical phenomena encountered at first hand in the solar system. Sunspots, solar flares, coronal mass ejections, the solar wind, collisionless shocks, magnetospheres, radiation belts, and auroras are just a few of the many phenomena that are unified by the common set of physical principles of plasma physics. These processes operate in other astrophysical systems as well, but because these systems can be examined only remotely, theoretical understanding of them depends to a significant degree on the knowledge gained in the studies of the local cosmos. This report, *Plasma Physics of the Local Cosmos,* by the Committee on Solar and Space Physics of the National Research Council's Space Studies Board attempts to define and systematize these universal aspects of the field of solar and space physics, which are applicable elsewhere in the universe where the action is only indirectly perceived.

The plasmas of interest to solar and space physicists are magnetized—threaded through with magnetic fields that are often "frozen" in the plasma. In many cases, the magnetic field plays an essential role in organizing the plasma. An example is the structuring of the Sun's corona by solar magnetic fields in a complex architecture of loops and arcades—as seen in the dramatic close-up views of the solar atmosphere provided by the Earth-orbiting TRACE observatory. In other cases, such as the Sun's convection zone, the plasma organizes the magnetic field. Indeed, it is the twisting and folding of the magnetic field by the motions of the plasma in the solar convection zone that amplifies and maintains the Sun's magnetic field. In all cases, however, the plasma and the magnetic field are intimately tied together and mutually affect each other. The theme of magnetic fields and their interaction with plasmas provides an overall framework for this report. An overview is presented in Chapter 1, introducing the chapters that follow, each of which treats a particular fundamental set of phenomena important for our understanding of solar system and astrophysical plasmas.

The question of how magnetic fields are generated, maintained, and amplified, together with the complementary question of how magnetic energy is dissipated in cosmic plasmas, is explored in the second chapter of this report, "Creation and Annihilation of Magnetic Fields." The focus is on the *dynamo* and on *magnetic reconnection*. Chapter 2 discusses the current understanding of the workings of these processes in both solar and planetary settings and identifies several outstanding problems. For example, understanding how the differential rotation of the solar interior arises represents a significant challenge for solar dynamo theory. In the case of planetary dynamos, important open questions concern the role of physical processes other than the Coriolis force in determining the morphology and alignment of the magnetic field (e.g., of Uranus and Neptune) and the influence of effects such as fluid inertia and viscous stress on Earth's dynamo. With respect to magnetic reconnection, a significant advance in our understanding has been achieved with the development of the kinetic picture of this process. However, what triggers and maintains the reconnection process is the subject of great debate. Moreover, how reconnection operates in three dimensions is not well understood.

Chapter 3, "Formation of Structures and Transients," examines some of the important structures that are found in magnetized plasmas. These include *collisionless shocks*, which develop when the relative velocity between different plasma regimes causes them to interact, producing sharp transition regions, and *current sheets*, which separate plasma regions whose magnetic fields differ in orientation and/or magnitude. A transient structure that occurs in a number of different plasma environments (solar active regions, the corona, the solar wind, the magnetotail) is the *flux rope*, a tube of twisted magnetic fields. Scientists have learned much about the plasma structures in our solar system but still have numerous questions. Studies of Earth's bow shock have provided basic understanding of shock dissipation and shock acceleration in collisionless plasmas, but much work remains in extending this understanding to large astrophysical shocks. This will require understanding of strong interplanetary shocks in the outer heliosphere and, ultimately, direct observation of the termination shock. Flux ropes have also been extensively observed, but many unanswered questions remain: How are flux ropes formed and how do they evolve? What determines their size? How are they destroyed? What is their relation to magnetic reconnection?

Chapter 3 also examines magnetohydrodynamic turbulence, a phenomenon that is a classic example of the way in which magnetized plasmas couple strongly across multiple spatial and temporal scales. In turbulent coupling, energy is fed into the largest scales and then progressively flows down to smaller scales, eventually reaching the "dissipation scale," where heating of the plasma occurs. Turbulence has been most completely studied in the solar wind, but questions remain concerning the detailed structure of heliospheric turbulence and how this structure affects energetic particle scattering and acceleration. Turbulent processes also occur in the Sun's chromosphere as well as in Earth's magnetopause and magnetotail. Outstanding problems include the role of turbulence in transport across boundary layers, the onset of turbulence in thin current sheets, and the coupling of micro-turbulence to large-scale disturbances.

Plasmas throughout the universe interact with solid bodies, gases, magnetic fields, electromagnetic radiation, and waves. These interactions can be very local or can take place over regions as large as the size of galaxies. Chapter 4 discusses four classes of plasma interaction. *Electromagnetic interaction* is exemplified by the coupling of a planetary ionosphere and magnetosphere by electrical currents aligned with the planet's magnetic field. The aurora is a familiar and dramatic manifestation of the energy transfer that results from this coupling. Electromagnetic coupling is also believed to be important in stellar formation, through the redistribution of angular momentum between the protostar and the surrounding nebular material. *Flow-object interactions* refer to the processes that occur when plasma flows past either a magnetized or an unmagnetized object. Typical processes include reconnection, turbulent wakes, convective flows, and pickup ions. The third class of plasma interactions are those that involve the *coupling of a plasma with a neutral gas,* such as the exchange of charge between ions and neutral atoms or collisions between ions and neutrals in Earth's auroral ionosphere, which drive strong thermospheric winds. The final category is *radiation-plasma interactions,* which is important for understanding the structure of the Sun's corona: radiation-plasma interactions produce a monotonically decreasing temperature-altitude profile in the corona in great contrast to a falling-then-rising profile produced by the standard quasi-static models.

Chapter 5, "Explosive Energy Conversion," treats the buildup of magnetic energy and its explosive release into heated and accelerated particles as observed in solar flares, coronal mass ejections, and magnetospheric substorms. Since the first observation of a solar flare in 1859 and the recognition that solar disturbances are associated with auroral displays and geomagnetic disturbances, magnetic energy release has been a central topic of solar-terrestrial studies. Because of their potentially disruptive influence on both ground-based and space-based technological systems, such explosive events are of practical concern as well as of great intrinsic scientific interest.

Both solar flares and coronal mass ejections (CMEs) result from the release of magnetic energy stored in the Sun's corona. It is not understood, however, how energy builds up and is stored in the corona or how it is then converted into heating in flares or kinetic energy in CMEs. At Earth, magnetic energy stored in the magnetotail through the interaction of the solar wind and the magnetosphere is explosively released in substorms, periodic disturbances that convert this energy into particle kinetic energy. The details of how stored magnetic energy is transferred from the lobes of the magnetotail to the plasma sheet and ultimately dissipated remain subjects of intense debate. The storage and release of magnetic energy occur universally in astrophysical plasmas, as evidenced by the enormous flares from M-dwarfs and the stellar eruption observed in the young XZ-Tauri AB binary system. What is learned about the workings of magnetic storage-release mechanisms in our solar system is likely to contribute to our understanding of analogous processes in other, remote astrophysical systems as well.

The key mechanisms by which magnetized plasmas accelerate charged particles are reviewed in Chapter 6, "Energetic Particle Acceleration." *Shock acceleration* occurs throughout the solar system, from shocks driven by solar flares and CMEs to planetary bow shocks and the termination shock near the boundary of the heliosphere. Particles are accelerated at shocks by a variety of mechanisms, and the resulting energies can be quite high, >100 MeV and even in the GeV range for solar energetic particles accelerated at CME-driven shocks. One topic of particular interest in current shock acceleration studies is the identity of the particles that form the seed population for the shock-accelerated ions. What, for example, are the sources and composition of the pickup ions that are accelerated at the termination shock to form anomalous cosmic rays?

Coherent electric field acceleration arises from electric fields aligned either perpendicular or parallel to the local magnetic field. Induced electric fields perpendicular to the geomagnetic field play a role in the radial transport and energization of charged particles in Earth's magnetosphere and contribute to the growth of the outer radiation belt during magnetic storms. Parallel electric fields accelerate auroral electrons and accelerate plasma from reconnection sites; they are also involved in the energization of solar flare particles. *Stochastic acceleration* results from randomly oriented electric field perturbations associated with magnetohydrodynamic waves or turbulence. It plays a role in the acceleration of particles in solar flares, in the acceleration of interstellar pickup ions in the heliosphere, and possibly in the acceleration of relativistic electrons during geomagnetic storms.

All of these acceleration mechanisms may occur simultaneously or at different times. For example, direct energization of particles by electric fields, interactions with ultralow-frequency waves, and localized, stochastic acceleration may all contribute to the storm-time enhancement of Earth's radiation belt. However, in this case as in others, distinguishing among the various acceleration mechanisms as well as determining the role and relative importance of each poses challenges to both the observational and the theory and modeling communities.

Plasma Physics of the Local Cosmos examines the universal properties of solar system plasmas and identifies a number of open questions illustrative of the major scientific issues expected to drive future research in solar and space physics. Recommendations regarding specific future research initiatives designed to address some of these issues are offered in another recent National Research Council report, *The Sun to the Earth—and Beyond: A Decadal Research Strategy for Solar and Space Physics*, which was prepared by the Solar and Space Physics Survey Committee under the auspices of the Committee on Solar and Space Physics.[1] The two reports are thus complementary. The Survey Committee's report presents a strategy for investigating plasma phenomena in a variety of solar system environments, from the Sun's corona to Jupiter's high-latitude magnetosphere, while *Plasma Physics of the Local Cosmos* describes the fundamental plasma physics common to all these environments and whose manifestations under differing boundary conditions are the focus of the observational, theoretical, and modeling initiatives recommended by the Survey Committee and its study panels.

NOTE

1. National Research Council, *The Sun to the Earth—and Beyond: A Decadal Research Strategy in Solar and Space Physics*, The National Academies Press, Washington, D.C., 2003. See also *The Sun to the Earth—and Beyond: Panel Reports*, 2003, the companion volume containing the reports of the five study panels that supported the survey.

1

Our Local Cosmic Laboratory

Plasma is the fourth state of matter and is ubiquitous in the universe. Plasmas pervade intergalactic space, interstellar space, interplanetary space, and the space environments of the planets. With the help of magnetic fields, plasma organizes itself into galactic jets, radio filaments, supernova bubbles, accretion disks, galactic winds, stellar winds, stellar coronas, sunspots, heliospheres, magnetospheres, and radiation belts. Magnetic fields partition space into tubes and shells of all sizes from galactic to planetary scales. Plasmas generate cosmic rays, stellar flares, coronal mass ejections, interstellar and interplanetary shock waves, magnetospheric storms, and a cacophony of radio waves. Plasmas absorb energy flowing steadily from the nuclear reactions within stars and from angular momentum shed by spinning magnetized bodies and release it explosively as x-rays and energetic particles. Structured, dynamic, and permeating apparently "empty" space, cosmic plasmas moderate energy flow across an enormous range of space and time scales.

Our local space environment—the heliosphere with its central star (the Sun) and orbiting planets—provides examples of many of the structures and processes that cosmic plasmas exhibit. Because of its accessibility to space probes, it is a local laboratory for in situ astrophysical plasma research. Eugene Parker has noted: "The little piece of cosmic real estate that we call our own, or can probe with spacecraft, is the most important corner of the universe for astronomical research."[1] The discipline of solar and space physics concentrates on understanding the local space environment. This report examines some of the universal properties of cosmic plasmas that have been identified from the unique knowledge base provided by nearly a half century of solar and space physics research. This general scientific understanding of the complex dynamics of magnetized plasmas forms the basis for extrapolation to remote astrophysical plasma systems, inaccessible to direct study.

From the perspective of pure science, plasma astrophysics offers the deep intellectual challenge of understanding the universe as a collection of self-organized, multiscale, coupled systems of space plasma structures and processes. Phenomena unpredictable by analytical theory emerge from such complex systems. For example, Richard Feynman notes: "Our equations for the sun . . . as a ball of hydrogen gas, describe a sun without sunspots, without the rice-grain structure of the surface, without prominences, without coronas."[2] Eugene Parker could predict the solar wind and the spiral magnetic field, but after decades of observations no one has predicted stellar flares or storms within magnetospheres. Without measurements within our local cosmic laboratory, we still would be oblivious of coronal mass ejections (CMEs), the most powerful local manifestations of cosmic storms. The CME epitomizes the dynamics of cosmic plasmas—a burst of energy on a global (heliospheric) scale drives convective (magnetospheric) motions on a macroscale. These motions, in turn, induce flow shears on a mesoscale (magnetotail) that stretch and stress magnetic fields that finally snap on a microscale (local reconnection) owing to instability. The snap initiates an explosion that triggers powerful energy release on every scale. Plasma processes throughout the universe are, by and large, variations on this theme.

During more than 40 years of progress marked by probes of geospace, visits to all our solar system planets but one and to six moons, three comets, and two asteroids, and spacecraft sailing to the edge of the heliosphere, the field of solar and space physics has observed and analyzed the many forms taken by magnetized plasma in the solar system. By documenting the particular attributes and behavior of solar system plasmas, the field of solar and space physics has been conducting fundamental plasma science within a unique natural laboratory—one in which plasma-physical phenomena can be studied in situ and without the limitations to which experiments in ground-based laboratories are subject. Sufficient knowledge has been amassed during the past four-plus decades that the study of fundamental plasma processes within our local cosmic laboratory is now considered an essential component of solar and space physics. By investigating these plasmas as they manifest themselves in the spacecraft-accessible regions of the solar system, we can explore and understand the structures and dynamics of magnetized plasmas throughout the more distant cosmos.

CONTRIBUTIONS TO UNDERSTANDING COSMIC PLASMAS

To illustrate the potential of solar and space physics to benefit other fields, this section recounts contributions that such studies have already made. The discovery in the second half of the 19th century of a phenomenon that we now call solar flares gave the first hint that cosmic plasmas have a propensity for explosive energy release. Since then, this tendency has revealed itself whenever instruments with new eyes have looked, making sudden energy release in the cosmos a central theme in space physics and plasma astrophysics. The deep mystery of how the Sun influences the geomagnetic field—an influence Lord Kelvin dismissed as "a mere coincidence" but Sir John Herschel lauded as presaging "a vast cosmical discovery such as nothing hitherto imagined can compare with"—led a century later to the prediction of the solar wind.[3] Confirmation and generalization to stellar winds soon followed.

Solar and space physics has given science the concept of magnetospheres and the first viable model of a magnetic dynamo that can generate planetary, stellar, and galactic magnetic fields. In less than 20 years, dedicated space physics missions and modeling brought the subject of collisionless shocks from an oxymoron to one of the most sophisticated examples of data-theory closure in science. Collisionless shock theory has been applied to the study of particle acceleration in both space and astrophysical plasma regimes, leading to a deep understanding of the way in which solar energetic particles and anomalous and galactic cosmic rays are accelerated.

The study of what happens when the solar wind encounters the local interstellar medium (LISM) has given rise to the concept of the heliosphere, the region of space dominated by the solar wind and the interplanetary magnetic field. Although spacecraft have yet to reach the boundaries of this region, remote sensing observations have detected radio emissions from just beyond the collisionless shock formed by the solar wind's encounter with the LISM and have revealed the existence of a "wall" of interstellar hydrogen just upstream of the heliosphere. Loosely speaking, as the LISM flows around the heliosphere, interstellar neutral hydrogen piles up, forming a wall-like structure at the nose of the heliosphere. The concept of such a wall of interstellar material now drives research programs to look for interstellar hydrogen walls around other stars, several of which have been reported.

Cosmic plasmas emit radio waves that furnish the means to detect these plasmas from Earth. Studies by space physicists of auroral kilometric radiation provide a terrestrial example of how the coupling of in situ observations and theory has led to a detailed understanding of the electron-cyclotron maser instability, a wonderfully efficient mechanism for moving energy from particle motions into radio waves. This theory is finding wide application in interpreting emissions from all magnetized outer planets (in particular, Jupiter), impulsive solar flares, binary stellar systems, and flare stars.

A last example of contributions by solar and space physics that have wide application is magnetic reconnection, perhaps the most universally invoked concept in studies of cosmic plasmas. The theory of magnetic reconnection has recently joined the ranks of long-standing, tough problems that are well on the way toward satisfactory solution. Cracking the problem entails identifying which mechanisms from a large field of candidates are important, and then understanding the coupling between disparate mechanisms that operate on widely separated spatial scales.

THE IMPORTANCE OF MAGNETIC FIELDS IN THE UNIVERSE

A key to understanding cosmic plasmas is the role that magnetic fields play in their dynamics and structure. Magnetic fields can act as a source of pressure and can interact with plasmas to cause expansion (e.g., stellar winds and jets). The presence of magnetic fields often causes the motion of the plasma to be turbulent (e.g., in the solar wind, galactic radio jets, and Earth's magnetotail). In magnetized plasmas, magnetic energy is often explosively converted into particle kinetic energy (e.g., stellar flares and magnetospheric substorms). In many plasma regimes, the magnetic fields structure and organize the plasma. Magnetically structured matter tends to define shells, tubes, and sheets (e.g., radiation belts, flux ropes, and current sheets). The solar system serves as a local laboratory for the study of such universal properties of astrophysical plasmas.

LOCAL PLASMA ASTROPHYSICS

Astronomy and astrophysics are sciences that have mature aspects (e.g., many objects observed in the optical regime) as well as discovery-mode aspects (e.g., observations in new wavelength regimes that reveal fundamentally new phenomena). Plasma astrophysics, as practiced in the local solar system laboratory, that is, space plasma physics, is relatively mature. As a science, space plasma physics is moving beyond the initial discovery phase to one in which detailed understanding of the physics is being sought.

Much of what we have learned about the behavior of plasmas in space can be thematically organized in the following universal categories:

1. Creation and annihilation of magnetic fields,
2. Formation of structures and transients,
3. Plasma interactions,
4. Explosive energy conversion, and
5. Energetic particle acceleration.

These categories form the basis for the discussion in the chapters that follow. Figure 1.1 shows these topics and their contents as far as researchers have identified them.

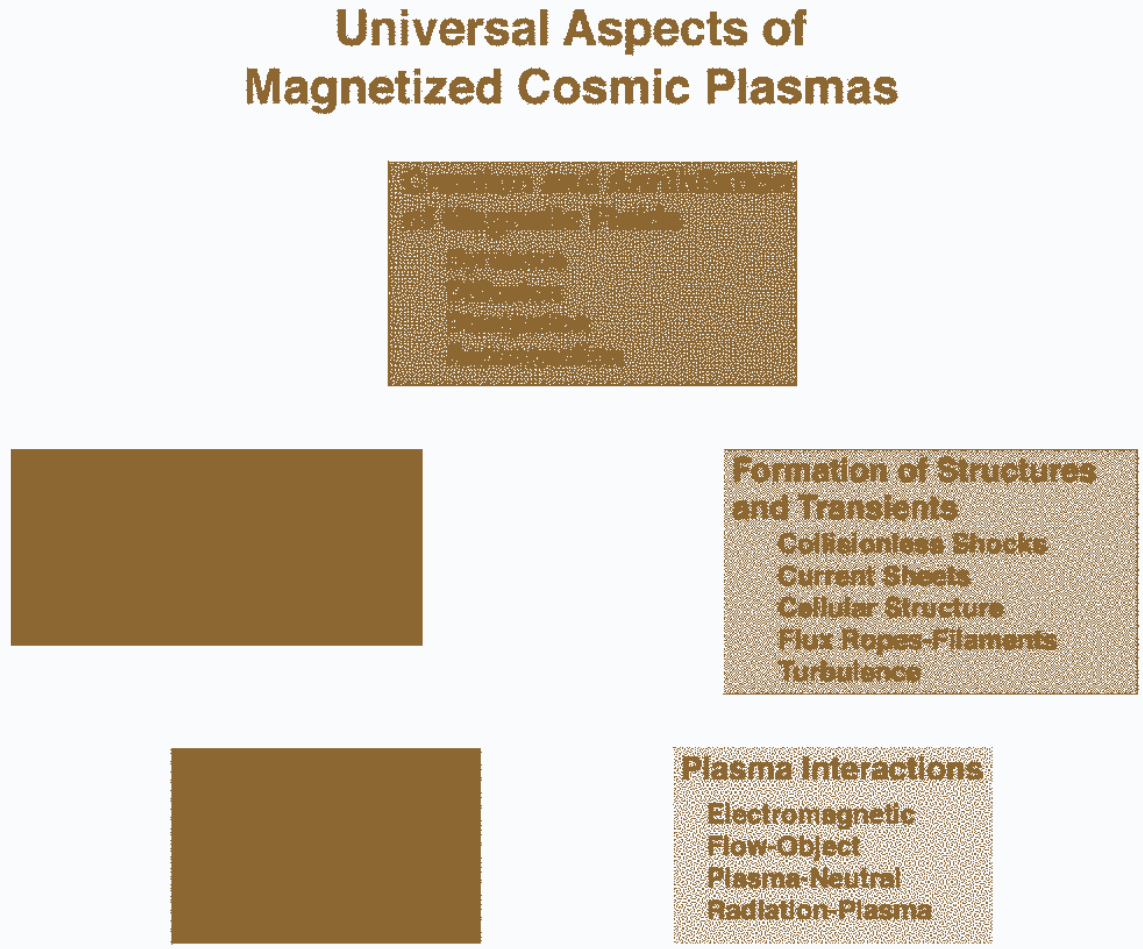


FIGURE 1.1 Five fundamental behaviors characteristic of magnetized cosmic plasmas.

The top box in Figure 1.1 is "Creation and Annihilation of Magnetic Fields." Cosmic magnetic fields result from an ever-evolving competition between creation by magnetic dynamos and destruction involving one or more of the following processes: diffusion, dissipation, and magnetic reconnection. Dynamos are evident on the Sun and within most planets (Mercury, Earth, evidently early Mars, and the giant planets) and within at least one moon (Ganymede). With respect to annihilation, magnetic reconnection deserves special mention because it is universal in two senses. First, it likely occurs wherever dynamos create magnetic fields—almost everywhere in the universe. Second, magnetic reconnection plays a central role in solar flares, coronal mass ejections, and the dynamics of magnetospheres.

Next in Figure 1.1 (moving clockwise) is the category "Formation of Structures and Transients." Collisionless shocks are ubiquitous in cosmic plasmas (e.g., planetary bow shocks, CME-driven interplanetary shocks, interstellar shocks associated with supernova remnants) and are important sites of particle acceleration. Shocks are created when the relative velocity between plasma regimes creates sharp transitions. Magnetism in plasmas spontaneously generates current sheets (e.g., the heliospheric current sheet and the magnetotail current sheet), cellular structures (e.g., coronal arcades and magnetospheres), flux ropes or filaments (e.g., plasmoids and sunspots), and turbulence (e.g., solar wind fluctuations and bursty bulk flows). The generation of filaments and flux ropes results from differential flows that stretch magnetic fields, which then, through instability or reconnection, segregate into coherent tubes of fixed flux. Current sheets spontaneously form whenever and wherever magnetized plasmas of different origins meet. They also spontaneously form when random velocity fields shuffle and twist field lines (such as in the Sun's photosphere).

Next in the circuit of Figure 1.1 is the category "Plasma Interactions." Plasmas interact with other plasmas and also with matter not in the plasma state. The solar wind interacts with planetary magnetospheres as well as with the ionospheres and neutral atmospheres of unmagnetized bodies such as Venus and comets. Planetary ionospheres and magnetospheres interact, with important consequences for both plasma regimes, as a result of their coupling by magnetic-field-aligned currents. Ionospheric plasmas interact collisionally with the neutral gases of planetary upper atmospheres, resulting in a mutual exchange of energy and momentum. Plasma interactions thus take a variety of forms and involve a number of different physical processes.

The next box in Figure 1.1 is "Explosive Energy Conversion," with examples of solar flares, CMEs, and substorms. The entry "solar flares" covers a hierarchy of phenomena from nanoflares, unresolvable by telescope, to importance-4, X-class bursts, visible to shielded but otherwise unaided eyes. The process called substorms at Earth appears to have analogues at Mercury and Jupiter. Explosive energy conversion occurs when magnetic energy builds slowly through stretching by differential flows and is released suddenly by one or more modes of instability. A key element is the role of magnetic reconnection in these processes—the merging of magnetic field lines is an efficient mechanism for generation of plasma flows and energy release. An important issue is whether differential flows that build magnetic energy or modes of instability that suddenly release it have properties in common. Is there a unified framework from which to understand explosive energy conversion as a manifestation of one or a few processes in different contexts? Or is each instance a case unto itself? This issue can be restated for nearly each example in Figure 1.1.

The remaining box in Figure 1.1 lists "Energetic Particle Acceleration" as a universal characteristic of magnetized plasmas. Solar system examples of energetic particle acceleration include anomalous cosmic rays, solar energetic particles, and radiation belts at Earth, Jupiter, Saturn, Uranus, and Neptune. The standing shocks of planetary magnetospheres, shocks associated with corotating interaction regions, and interplanetary shocks driven by CMEs all accelerate particles. The primary acceleration mechanism associated with shocks is known as Fermi acceleration, which results from the repeated passage of charged particles back and forth across the shock as they are reflected between the upstream and downstream plasma. Electric fields play a central role in the acceleration of charged particles in magnetized plasmas. These electric fields can be produced by time-varying magnetic fields (Faraday's law of magnetic induction), by charge-separation, and by the dissipation of Alfvén waves in planetary ionospheres. Coherent electric field acceleration is responsible, for example, for the acceleration of particles in solar flares, in Earth's magnetotail during magnetospheric disturbances, and in the auroral magnetosphere. Particle acceleration can also result from the action of plasma waves or turbulence (stochastic acceleration).

The intersection between space physics and plasma astrophysics provides fertile ground for the transfer of knowledge and generalization of specific, local cases to a much broader range of physical understanding of plasma processes in the universe.[4] As the chapters that follow demonstrate, there is a wide range of work that can now be used for continuing the evolution toward a closer relationship between space plasma physics and plasma astrophysics.

NOTES

1. Louis J. Lanzerotti, Charles F. Kennel, and E.N. Parker, eds., *Solar System Plasma Processes*, p. 378, North-Holland, New York, 1979.

2. R. Feynman, *Lectures*, Volume II, p. 41-12, Addison-Wesley, Boston, Mass., 1970.

3. On Lord Kelvin's skepticism and Herschel's enthusiasm, see E.W. Cliver, Solar activity and geomagnetic storms: The first 40 years, *Eos, Transactions, American Geophysical Union* 75(49), 569, 574-575, December 6, 1994; and Solar activity and geomagnetic storms: The corpuscular hypothesis, *Eos, Transactions, American Geophysical Union* 75(52), 609, 612-613, December 27, 1994.

4. On the intersection between space physics and plasma astrophysics, see also the chapter titled "Connections Between Solar and Space Physics and Other Disciplines" in the recent NRC report *The Sun to the Earth—and Beyond: A Decadal Research Strategy in Solar and Space Physics*, The National Academies Press, Washington, D.C., 2003.

2

Creation and Annihilation of Magnetic Fields

Magnetic fields exist throughout the universe, ranging from less than a micro-gauss in galactic clusters to 10^{12} gauss or more in the magnetospheres of neutron stars.[1] There is increasing evidence that these magnetic fields profoundly affect the fundamental dynamics of the universe through angular momentum transport during star formation, in the accretion of material onto stars and black holes, in the formation of jets, and in the creation of suprathermal gases responsible for much of the x-ray emission from a variety of astrophysical sources. Magnetic fields that are generated in astronomical bodies such as galaxies, stars, and planets produce forces that compete with convection and with rotational and gravitational forces. Within our own solar system the magnetic fields shed by the Sun interact with the fields surrounding Earth to produce the complex dynamics of the magnetosphere.

Because of the broad importance of magnetic fields in large-scale plasma dynamics, developing a first-principles understanding of the physical mechanisms that control the generation and dissipation of magnetic fields is an essential scientific goal. Magnetic fields are generated by the convective motions of conducting materials—plasma in most of the universe and conducting liquids in the case of planetary objects. The twisting and folding of the magnetic field by the motion of the conducting material lead to amplification of the field in a process known as the dynamo. Ultimately the growth of the magnetic field by the dynamo is limited by the field's back reaction on the fluid convection and by the dissipation of the magnetic energy. Thus, knowledge of the mechanisms by which magnetic fields are dissipated is essential to describing the overall amplification/saturation process of the magnetic fields.

The release of magnetic energy is often observed to occur in bursts, in essentially explosive processes that produce intense plasma heating, high-speed flows, and fast particles. Solar and stellar flares and magnetospheric substorms are examples of such explosive phenomena. Magnetic reconnection, in which oppositely directed magnetic field components rapidly merge to release the stored magnetic energy, has been identified as the dominant mechanism for dissipating magnetic energy. The description of the reconnection process is complicated by the need to describe correctly the small-scale spatial regions where the magnetic field lines change their topology. Surprisingly, kinetic effects at these very small scales have been found to strongly influence the release of magnetic energy over very large spatial scales.

This chapter briefly reviews the theoretical explanations that have been put forward for the creation of cosmic magnetic fields (the dynamo) and their annihilation (magnetic reconnection) and examines the operation of these processes in both solar and planetary settings.

MAGNETIC FIELD CREATION: DYNAMO THEORY

Many astrophysical bodies, including galaxies, stars, and planets, have an internally generated magnetic field. Although these bodies differ significantly in many aspects, they all possess within their interiors an electrically conducting fluid that is dominated by the Coriolis force because of their rapid rotation. In the case of the planets, the release of thermal and gravitational energy leads to convection in the planetary cores. In the case of stars and the Sun, convection is driven by heat from thermonuclear fusion. In many astronomical bodies the mean fields generated by the dynamo periodically reverse in time. A prominent example is the 22-year periodicity of the magnetic field of the Sun. To answer the question of the origin of magnetic fields, it is necessary to understand how magnetic fields are generated and maintained in rapidly rotating, convective fluids. This understanding is the goal of dynamo theory.

The dynamo process can be simply described as follows: a moving electrically conducting fluid stretches, twists, and folds the magnetic field. Dynamo action occurs if a small-amplitude seed magnetic field is sustained and amplified by the flow. The magnetic field increases in strength until the resultant magnetic forces are sufficient to feed back on the flow field. Dynamos can be quite complicated, and fundamental questions can be posed. How does a given flow generate a magnetic field? How does the generated magnetic field act to modify the flow? What energy source sustains the flow? While the first two questions can be studied within the context of magnetohydrodynamics, the answer to the last question depends on the specific physical system being studied. Finally, magnetic reconnection (in the generic sense of a mechanism that alters magnetic field topology) is an intrinsic part of any dynamo mechanism. The various magnetic field components that are generated by plasma flows must ultimately decouple and condense into a large-scale field (usually the dipole field in astronomical objects). The connectivity of field lines must change for this condensation to take place, which requires reconnection. What, therefore, are the processes that control magnetic reconnection in environments where dynamo action is important (e.g., the convection zone in the Sun or in the interior of planetary bodies)? In a self-consistent dynamo model, all these questions are related and so must be studied together.

Kinematic dynamo theory studies the generation of a magnetic field by a given flow. The importance of flow is described by the (nondimensional) magnetic Reynolds number R_m, defined as the ratio of magnetic diffusion time to the flow convection time. Dynamo action occurs if the growth rate of magnetic field perturbations is positive, that is, if the amplitude of an initially small perturbation increases with time. From kinematic theory the necessary condition for dynamo action is typically $R_m \geq 10$. The physical significance of this condition is that the electromotive force associated with the flow has to overcome the magnetic dissipation in the fluid in order for a dynamo to occur. Another important result of kinematic dynamo studies is the demonstration that an axisymmetric magnetic field cannot be generated by an axisymmetric flow. This result implies that dynamo action must be three-dimensional.

When the magnetic Reynolds number R_m is large (i.e., indicates a faster flow, or less electrical resistivity in the fluid), the field lines are "frozen" in to the flow and are thus stretched, twisted, and bent (Figure 2.1). In order for the net flux to increase, the field lines must reconnect (alter their topology). Because magnetic diffusion is weak, field line reconnection takes place in regions of small spatial scale. Overall, the dynamo process generates new magnetic field lines and the magnetic flux increases with time. A major mystery is the source of magnetic diffusion required to change the field topology, which greatly exceeds that resulting from classical collisional processes.

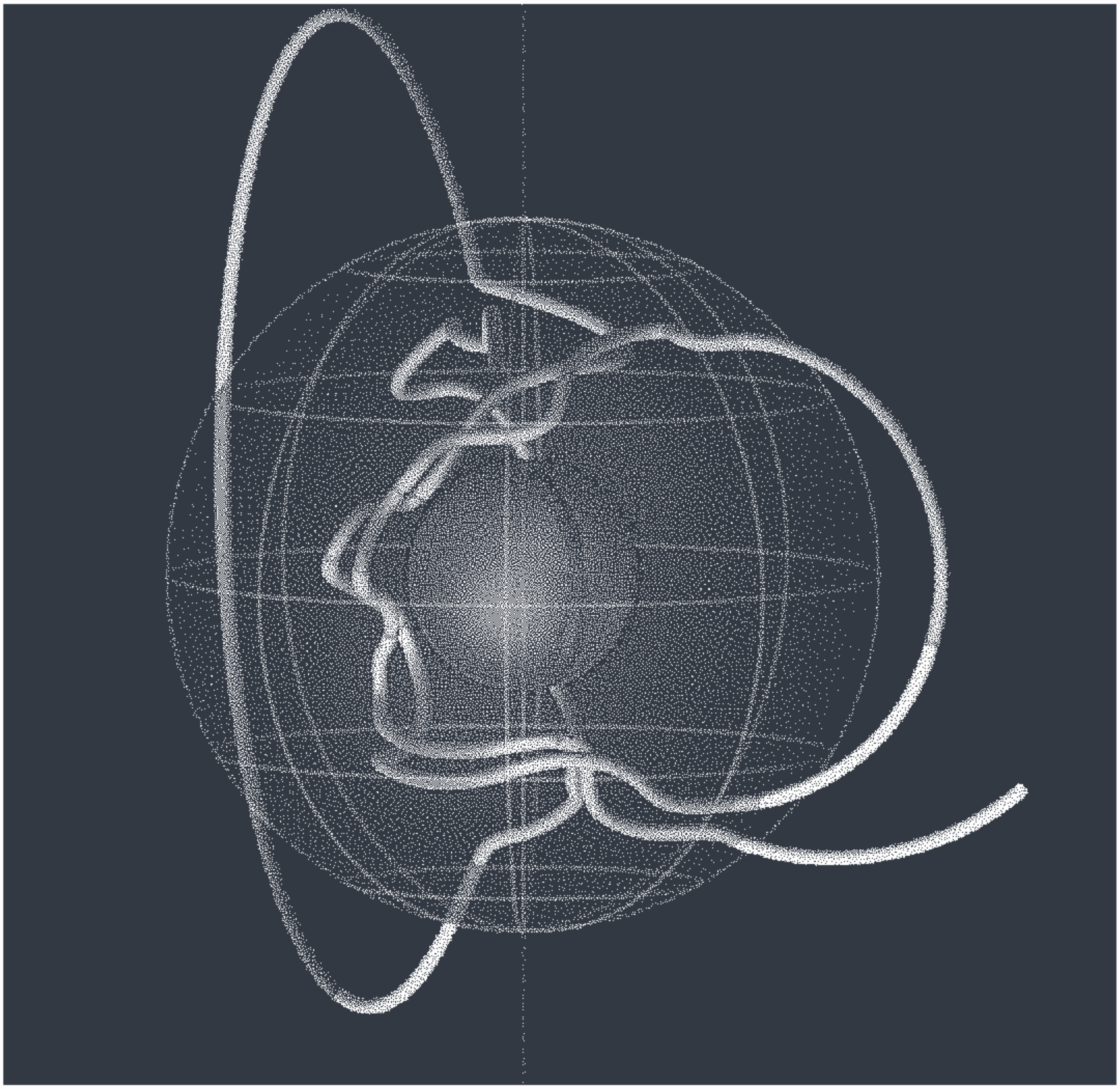


FIGURE 2.1 The stretching and twisting of a magnetic field line by fluid motion in Earth's outer core. Dynamo action occurs in the spherical shell between the outer blue surface, which represents the core-mantle boundary, and the red inner sphere, which represents the inner core. The yellow (blue) line segments in the figure indicate that the field line has a positive (negative) radial component. The field line is stretched in longitudinal directions by (zonal) differential rotations in the fluid core (the so-called ω-effect in dynamo theory) and is twisted in meridional directions by the cyclonic upwelling/downwelling flows (the so-called α-effect). Image courtesy of J. Bloxham (Harvard University). Reprinted, with permission, from W. Kuang and J. Bloxham, A numerical dynamo model in an Earth-like dynamical regime, *Nature* 389, 371-374, 1997. Copyright 1997, Macmillan Publishers Ltd.

For a given flow, there exists a critical value of R_m, at which the growth rate of the magnetic field perturbation is the largest. As R_m increases further, the growth rate of the large-scale magnetic fields decreases to zero, implying that a finite magnetic diffusivity (finite conductivity) of the fluid is necessary for dynamo action. This type of dynamo is often called a slow dynamo, to which class most models of Earth's dynamo belong.[2] However, kinematic dynamo studies also show that, for some three-dimensional chaotic flows, the growth rate of the large-scale magnetic field remains positive for large R_m. That is, dynamo action exists in the limit of vanishing magnetic diffusivity. This type of dynamo action is called a fast dynamo. For both cases it is essential that self-generation of the magnetic field occurs at spatial scales comparable in size to the entire region in which convection is taking place (e.g., the dipole field of the Sun or planets). That this is possible in the case of the fast dynamo has not been demonstrated.

While kinematic dynamo theory can well explain how a given flow generates a magnetic field, it does not take into account the influence of the generated magnetic field on the flow. The magnetic field lines do not passively follow the flow. They behave more or less like elastic threads. Therefore, in the process of stretching and bending the magnetic field lines, the flow also experiences a reaction force from the magnetic field. This magnetic force is called the Lorentz force and is proportional to the current density and the magnetic field in the fluid. The importance of the reaction forces can be assessed by comparing them to the leading-order forces (such as the Coriolis force in a rapidly rotating fluid like Earth's fluid core) in the fluid momentum equation.

CREATION OF MAGNETIC FIELDS IN THE SUN

Solar magnetic energy is continually being created, annihilated, and ejected. The physics underlying these opposing processes is known only in the most general terms, and detailed understanding faces significant theoretical and observational challenges. For example, although the Sun is the nearest star and the only star whose surface features can be resolved, much of the important action takes place on scales too small to be seen with existing telescopes. Telescopes detect the existence of the small-scale magnetic fields and motions but lack sufficient resolution to determine precisely what is happening. That important step must await the exploitation of adaptive optics on a telescope of large aperture.

The explosive dynamics observed in the atmosphere of the Sun originates in the gentle overturning of the gas in the convection zone, which occupies the outer 2/7 of the solar radius (1 solar radius = 7×10^5 km). The thermal energy in the central regions of the Sun diffuses outward as thermal black body radiation, with the temperature decreasing from 1.5×10^7 K in the central core to 2×10^6 K at the boundary between the radiative interior and the convection zone. Here, convective mixing takes over from radiative transport and delivers heat to the Sun's photosphere or visible surface. In addition to transporting thermal energy, the convection of the hot ionized (and hence electrically conducting) gas transports magnetic fields as well. The magnetic fields carried in the convection are stretched and contorted, with substantial increase in the magnetic energy. The magnetic fields are buoyant because they provide pressure without significant weight, and so they tend to bulge upward through the visible surface into the tenuous atmosphere above. Thus, they form the conspicuous bipolar magnetic regions that spawn sunspots, coronal mass ejections, and flares.

The hydrodynamics of the rotation of the Sun is described by the Navier-Stokes momentum equation, the equation for conservation of mass, the heat flow equation, and the ideal gas law. This model should reproduce the observed nonuniform rotation of the Sun and the meridional circulation, because both must be driven by the convection or they would have died out long ago as a consequence of the magnetic stresses. So far, however, this theoretical goal has not been achieved. Helioseismology has succeeded in mapping the internal rotation of the Sun, with the remarkable and unanticipated discovery that the radiative interior rotates approximately rigidly with a period of about 28 days, while the rotation of the convective zone varies with latitude but only weakly with depth. Thus, the observed surface rotation approximately projects downward to the base of the convective zone. The convective zone rotates with a period of 25 days at the equator, creating a strong forward shear where it meets the radiative zone. The rate of rotation decreases with increasing latitude, providing a period in the neighborhood of 35 days in the polar regions and creating a strong backward shear where the convective zone meets the radiative zone. Understanding how the differential rotation revealed by helioseismology arises represents a significant theoretical challenge.

The strong shear layer at the interface between the convective and radiative zones is known as the tachocline.[3] Rotational shear in this region plays a major role in the operation of the solar dynamo. The generation of the solar magnetic field involves the production of an azimuthal field from an initial poloidal field and the subsequent regeneration and amplification of the poloidal field from this azimuthal field by cyclonic convection. The nonuniform toroidal rotation shears the poloidal field, producing an azimuthal magnetic field. An individual cyclonic convective cell creates an upward bulge (an Ω loop) in the azimuthal field, which it rotates into the meridional plane (Figure 2.2). The result of the generation of many such loops, after smoothing by diffusion, is the development of a mean magnetic field in the meridional plane, thereby supplementing the original poloidal field. These processes are described by the magnetohydrodynamic dynamo equations, first written down 50 years ago.

The solutions of the magnetohydrodynamic dynamo equations in the convective zone of the Sun are periodic with a time scale of around 22 years, resembling the observed periodicity of the Sun's magnetic field.[4] However, there is still much to be understood. For example, the inferred "turbulent" diffusion of the magnetic field in the convective zone, which is essential in establishing the proper scale and period of the solar magnetic field, is not understood. In addition, the magnetic fields extending through the visible surface of the Sun actually consist of unresolved, widely spaced, very intense (1500 gauss) flux bundles (fibrils) with diameters around 100 km. Measurements from the TRACE satellite suggest that these magnetic fields form a dense and dynamic layer of magnetic loops in the corona, dubbed a "magnetic carpet."

Outstanding Questions About the Creation of Solar Magnetic Fields

- What is the physical mechanism for the diffusion of strong magnetic fields in the Sun?
- Why does the magnetic field at the surface of the Sun take the form of bundles of flux or fibrils?
- What produces the differential rotation as a function of radius and latitude that helioseismology has revealed in the Sun's interior?
- What causes the approximate 22-year magnetic cycle and why do its strength and period vary over the centuries?

PLANETARY DYNAMOS

Like the Sun, many planets self-generate, or at one time self-generated, magnetic fields. The existence of a terrestrial magnetic field was established some four centuries ago, although it was mistakenly attributed to a mass of permanently magnetized material in Earth's interior. In the past few decades, NASA space missions have discovered internal magnetic fields at five other planets—Mercury, Jupiter, Saturn, Uranus, and Neptune—and at the jovian moon Ganymede. Moreover, the recent discovery of a strong crustal magnetic field at the surface of Mars by the Mars Global Surveyor suggests that that planet, too, once possessed a strong internal magnetic field. The general principles of dynamo action in rotating, convecting, electrically conducting fluids are much the same in the Sun and the planets. However, the specific

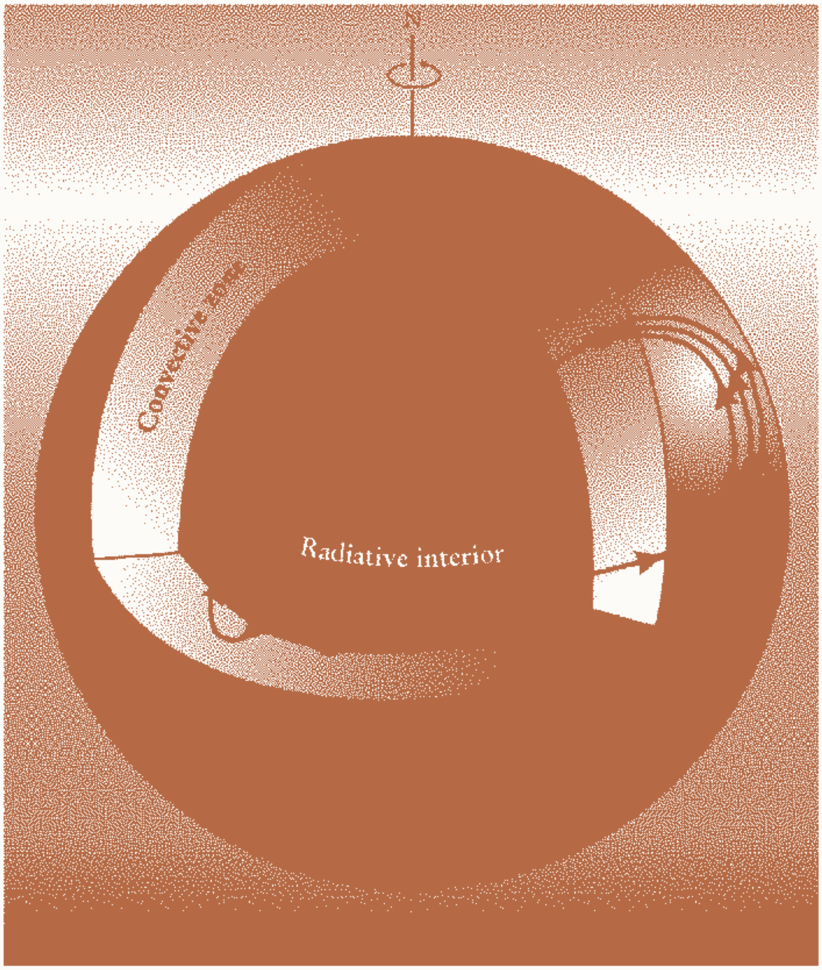


FIGURE 2.2 Schematic illustrating the interplay of rotational and cyclonic-convective forces in the operation of the solar dynamo. Strong toroidal or azimuthal fields are generated from an existing poloidal field in the tachocline, a region of strong shear at the base of the convection zone. Cyclonic convection pushes a bulge in the azimuthal field and rotates it into the meridional plane. Image courtesy of E. Plotkin (American Institute of Physics). Reprinted, with permission, from E.N. Parker, The physics of the Sun and the gateway to the stars, *Physics Today* 53(6), 26-31, June 2000. Copyright 2000, American Institute of Physics.

---

conditions are sufficiently different that planetary dynamos are a subject unto themselves. While the magnetic fields of the Sun are generated near its surface, in the terrestrial planets the dynamo is confined to the planetary core, which is shielded from the atmosphere by the crust and mantle. The giant planets approach more closely the solar case, with the convection zone extending to the planetary surface.

Observational and theoretical studies of planetary magnetic fields began with the study of the geomagnetic field. Applications associated with the geomagnetic field date back to the first century A.D. (e.g., the invention of the compass). But the first serious study of the origin of the geomagnetic field appeared much later, following William Gilbert's proposal in *De Magnete* (1600) that Earth is a great magnet. Later Karl

Friedrich Gauss provided the mathematical tools to separate the internal magnetic field from the external magnetic field. They are still used in today's analyses of the geomagnetic and planetary magnetic fields. In the 1940s, Walter Elsasser initiated the development of hydromagnetic dynamo theory, which is the basis for our understanding of the geomagnetic field and of internally generated planetary fields.

The development of planetary dynamos is closely correlated with the thermal evolution of the planets. A simple picture is that, at the accretion of the planets, tremendous gravitational energy was transferred into thermal energy, resulting in the formation of molten, electrically conducting planetary cores. As the planets cool off (e.g., the secular cooling of Earth), heat is released from the planetary interiors. Convection in the planetary cores facilitates the fast cooling rate, and the convective flows drive the internal dynamo. Other possible energy sources for the dynamo have also been proposed, such as radiogenic heat and tidal force; the latter source is still being considered in geodynamo studies.

The best-studied convection-driven planetary dynamo is that of Earth. Earth possesses a large fluid outer core, with a radius of approximately 3200 km, which is about half Earth's mean radius, and a solid inner core with a radius of 1100 km. The molten alloy in Earth's outer core is iron-rich (and thus electrically conducting), with smaller amounts of lighter constituents (e.g., oxygen, sulfur). In the secular cooling process, the inner core grows outward because of the freezing of the liquid iron at the inner-core boundary. The lighter constituents and latent heat are thus released into the outer core, producing strong buoyancy forces that drive the convection that is necessary for the geodynamo.

Mercury is the only other terrestrial planet to possess a strong intrinsic magnetic field today. Mercury's field, the existence of which was revealed by Mariner 10 observations in the mid-1970s, is generally thought to be generated by dynamo action in a fluid outer core. However, questions remain about whether the present-day existence of a partially molten core is consistent with Mercury's thermal history, and alternatives to a hydromagnetic dynamo have been proposed. In the case of Mars, which today possesses no, or only a very weak, intrinsic field, theoretical studies suggest that the cooling rate (and thus the buoyancy force) was sufficient to drive an internal dynamo only during the first 100 million to 150 million years of the planet's history. Mars's remanent crustal magnetic field has been mapped by the Mars Global Surveyor. The imprints of the internal field in the crust reveal the history of the martian dynamo and may provide evidence of variations in the thermal processes that occurred in the martian mantle. Venus, like Mars, has no apparent intrinsic field, but unlike the case with Mars there is insufficient evidence about a possible crustal field to support conclusions about the existence of an internal dynamo at an earlier stage in Venus's evolution.

The dynamos of the outer planets operate in planetary interiors quite different from those of the terrestrial planets. While convection in these planets may extend to the surface, dynamo action occurs in metallic hydrogen (Jupiter and Saturn) or ionic (Uranus and Neptune) cores. Most of the internal field and perhaps the surface flow could in principle be measured, thus permitting more direct observation of the dynamo action.

The recent numerical modeling of planetary dynamos has been very successful and is rapidly becoming the main tool for studying in detail the nonlinear dynamics of dynamo action. Although the mathematical models are very simple compared to the actual planetary cores, they can produce solutions that agree qualitatively with observations. In particular, geodynamo modeling has shown that a predominantly dipolar magnetic field exists at the core-mantle boundary.[5] The westward drift of the modeled geomagnetic field is comparable to that inferred from geomagnetic observations. Numerical simulations also demonstrate repeated reversals of the polarity of the magnetic field, a phenomenon that is well known from the paleomagnetic records.

Despite much progress in studies of planetary dynamos, many long-standing fundamental problems remain unanswered, while the results of numerical dynamo modeling have given rise to new questions. The dominance of the Coriolis force is invoked to explain the nearly axisymmetric dipolar geomagnetic field—

that is, to account for the fact that the magnetic dipole axis is very close to the rotation axis. This explanation cannot be generalized to dynamo action in all rapidly rotating fluids. For example, observations have revealed a very different magnetic field geometry at Uranus and Neptune: The field structures of both planets have no obvious correlation with the rotation axis, suggesting that other physical processes in the dynamo must be important in determining the morphology of the magnetic field. Recent studies have focused mainly on the effect of the geometry of the fluid core on the generated magnetic field. However, the strength of the driving force could also be important.

Although numerical geodynamo modeling has been a great success, the relative roles of the dominant forces inside Earth's core are still not well understood. In Earth's core (as a rapidly rotating fluid with a strong field dynamo), the Coriolis force, the buoyancy force (the driving force for convection), and the Lorentz force are the leading-order forces in the momentum balance of the flow. Fluid inertia and viscous stress are very small and are neglected in the leading-order approximation. However, numerical modeling shows that variations of these higher-order effects could lead to very different dynamical processes inside the core, although the generated magnetic fields are similar at the core-mantle boundary. Further study of the dominant forces acting in Earth's core (and in general, in a rapidly rotating fluid) is therefore necessary. Observations of other physical quantities of Earth, such as the gravity field and surface deformation, may help in identifying the dynamical processes that are most active in the core.

Outstanding Questions About Planetary Dynamos

- What is the dependence of the dynamo on the properties of the planetary interior—in particular, on the various dissipative parameters of the conducting fluids?
- Besides the Coriolis force, what are the physical processes in the dynamo that determine the configuration (including the alignment) of planetary magnetic fields?
- What are the turbulent flow structures in planetary cores?

MAGNETIC FIELD ANNIHILATION: RECONNECTION THEORY

A variety of phenomena in the universe are powered by the sudden release of magnetic energy and its conversion into heat and high-velocity plasma flows. Understanding such phenomena, and therefore the mechanism by which magnetic energy is released, has occupied space physicists, astrophysicists, and plasma physicists for nearly five decades. Energy release rates calculated on the basis of classical ohmic or resistive dissipation are orders of magnitude too small to explain the observed time scales on which stored magnetic energy is released in events such as solar flares. A more efficient mechanism for magnetic energy release is therefore required. (Ohmic dissipation rates can be characterized by the resistive dissipation time $\tau_r = 4\pi L^2/\eta c^2$, which is the time required for the energy in a system with scale size L with resistivity η to dissipate a significant fraction of the magnetic energy.)

Scientists very early on proposed magnetic reconnection as the mechanism by which magnetic energy could be released on a much shorter time scale than is possible through simple resistive dissipation.[6] The reconnection process is illustrated in Figure 2.3, which shows the results of a kinetic simulation (particle ions and fluid electrons). In the top panel oppositely directed magnetic field lines "reconnect" to form a topological x-line configuration. The resulting bent field lines attempt to straighten out and in doing so drive high-speed flows outward from the x-line as shown in the second panel. The outward flows produce a pressure drop in the vicinity of the x-line that draws in regions of reversed magnetic field toward the x-line. The entire process is therefore self-sustaining. The characteristic velocity associated with the outward flows is the Alfvén velocity, $v_A = \mathbf{B}/(4\pi\rho)^{1/2}$, where **B** is the magnetic field strength and ρ is the plasma mass

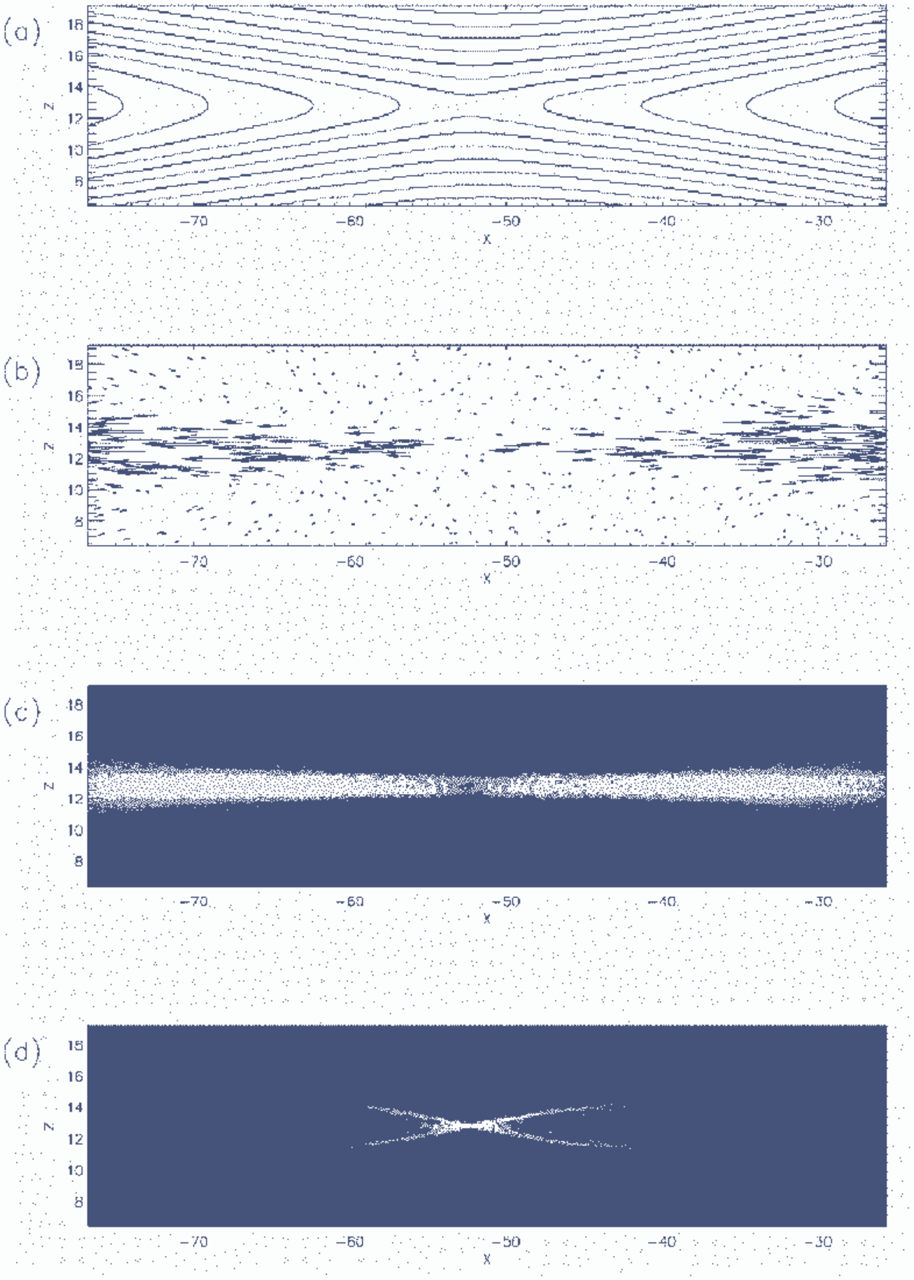


FIGURE 2.3 The reconnection of oppositely directed magnetic field components occurs within a spatially limited region known as the dissipation or diffusion region. Shown are the results of a hybrid simulation (particle ions and fluid electrons) of this region. In panel (a) oppositely directed magnetic fields "reconnect" to form a magnetic x-line configuration. The bent fields to the left and right of the x-line act like oppositely directed "slingshots" that expand outward to release their energy, driving the high-speed outflows shown in panel (b). The out-of-plane currents of ions and electrons in (c) and (d) sustain the magnetic configuration in (a). The distinct scale lengths of the ion and electron currents indicate that the motion of the two species has decoupled. In recent theoretical models the decoupling of electron and ion motion in the dissipation region is essential to achieving the fast reconnection observed in nature. Reprinted, with permission, from M.A. Shay et al., The scaling of collisionless, magnetic reconnection for large systems, *Geophysical Research Letters* 26(14), 2163-2166, 1999. Copyright 1999, American Geophysical Union.

density. Thus, a new time scale, the Alfvén time $\tau_A = L/v_A$, is important in magnetic reconnection. This time scale is shorter than the measured energy release times.

The rate of magnetic reconnection depends ultimately on the mechanism by which oppositely directed field lines reconnect. In an ideal plasma with no dissipation, the magnetic field is "frozen" in the plasma. That means that no topological change in the magnetic field is possible. Dissipation must therefore play a role in facilitating the reconnection process. In order for the intrinsically weak dissipative process to compete with Alfvénic flows, the dissipation must occur at small spatial scales. The scientific challenge has therefore been to develop models of the very localized dissipation or diffusion region that develops around the x-line to facilitate the topological change in magnetic field required for reconnection to occur.

P.A. Sweet and E.N. Parker, who developed the earliest model of the dissipation region, explored the dynamics of the thin current layer separating two macroscopic regions of an oppositely directed magnetic field. The resultant energy release time is given by the hybrid of the resistive and Alfvén times, $(\tau_A \tau_r)^{1/2}$. Based on classical resistivity, this release time remains far too long to explain the observations. The narrow dissipation region of the Sweet-Parker model acts as an effective nozzle that severely limits the inflow velocity into the x-line. Enhanced resistivity, resulting from the turbulence associated with instabilities generated by the intense currents produced in the dissipation region, has often been invoked to shorten the energy release times. However, a solid theoretical foundation for such "anomalous resistivity" has been lacking.

H.E. Petschek and subsequent authors proposed that, if slow shocks formed at the boundary between the inflow and outflow regions, the length of the dissipation region could be shortened, allowing the outflow region to open up and therefore enhancing the rate of reconnection. One effect of the slow shocks would be to accelerate the inflowing plasma up to the Alfvén velocity of the outflow. Theoretical energy release times as short as the Alfvén time multiplied by logarithmic factors of the resistivity rendered reconnection rates fast enough to explain the observations even with very small values of classical resistivity. Simulations, however, have supported the Sweet-Parker rather than the Petschek picture. Simulations with a simple, constant but low resistivity produced dissipation regions with a macroscopic extent along the outflow, consistent with the Sweet-Parker model and therefore with slow reconnection. Models with enhanced resistivity in regions of high current were required to produce fast Petschek reconnection.

The Sweet-Parker and Petschek models address the problem of reconnection in terms of magnetohydrodynamic (MHD) theory. Recent research has emphasized the importance of kinetic (non-MHD) effects in facilitating reconnection and has employed numerical simulations and analytical theory to explore such effects.[7] The inclusion of kinetic effects has proven essential to understanding magnetic reconnection in Earth's and planetary magnetospheres, where classical collisions are negligible. The kinetic model has also arguably proven essential to efforts to understand reconnection in the solar atmosphere and possibly in the broader astrophysical context as well.

The results of the hybrid (particle ions and fluid electrons) simulation of reconnection shown in Figure 2.3 illustrate the multiscale structure of the dissipation region. At large scales, electrons and ions move together toward the x-line, where the change in magnetic topology occurs. Close to the x-line the ions decouple from the magnetic field and from the electrons, while even closer the electrons also decouple from the magnetic field. As a consequence, the out-of-plane ion and electron currents shown in panels (c) and (d) of Figure 2.3 have distinct spatial scales. Because of their greater mass, the unmagnetized ions occupy a much larger region than that occupied by the unmagnetized electrons. The key point is that the dynamics of the dissipation region where ions are unmagnetized is controlled by a class of dispersive waves (whistler or kinetic Alfvén waves) rather than by the usual magnetohydrodynamic Alfvén waves. Outside the small region close to the x-line, the resulting flow patterns closely mirror those of the Petschek model (no evidence for a macroscale Sweet-Parker current sheet), and the rates of reconnection are a substantial fraction of the Alfvén speed. It is the very high speed electron flows generated by these dispersive waves close to the x-line that remarkably facilitate fast energy release in a macroscopic system.

Such fast rates of reconnection appear to be consistent with solar, magnetospheric, and laboratory observations. The generation of dispersive waves at small scales is apparently the key to understanding fast reconnection as observed in nature. The benchmarking of this kinetic model with observations has been challenging for two reasons. First, it has only been in the past couple of years that the essentials of the kinetic models have emerged. Second, the small scale size of the dissipation region (of the order of tens of meters in the solar atmosphere and tens of kilometers in the magnetosphere) makes the acquisition of data very difficult. Nonetheless, data from recent satellite missions are for the first time beginning to document and confirm the essence of the kinetic reconnection model. More direct comparison between observations and theoretical models will be required to demonstrate that the theory correctly describes processes occurring in nature and the laboratory.

The explosive release of energy associated with reconnection is consistent with inflow rates into the x-line at a significant fraction (0.01-0.1) of the Alfvén speed. What triggers the reconnection process, however, has been a subject of great debate. In laboratory tokamak experiments, for example, there are unresolved questions concerning the onset of the "sawtooth crash," in which energy is expelled from the core of the confined plasma as a result of reconnection. The onset condition for solar flares and coronal mass ejections is similarly poorly understood. Is the trigger linked to kinetic effects associated with the structure of the dissipation region, or is it a consequence of the global configuration of the system? If the latter explanation is correct, then why do all of the observable systems exhibit trigger phenomena? Further discussion of this issue appears in Chapter 5.

In the interest of clarity the committee has up to this point focused exclusively on a picture of reconnection expected for a two-dimensional system. There is, however, substantial observational evidence that the release of magnetic energy in nature either takes place in intrinsically three-dimensional magnetic configurations or develops three-dimensional structure as a result of the reconnection process. The data from the TRACE observations of the solar corona provide graphic evidence for the release of energy in three-dimensional loops. High-speed flows measured in Earth's magnetotail, which are believed to be driven by magnetic reconnection, are spatially localized in the plane perpendicular to the flow, indicating that reconnection does not occur at extended x-lines but rather in spatially localized regions. Intrinsically three-dimensional reconnection is therefore a topic of great importance, but one of which current understanding remains limited.

MAGNETIC RECONNECTION IN THE SUN'S CORONA

Magnetic field annihilation in the solar atmosphere typically proceeds in an explosive manner, producing flare energy releases over a broad range from 10^{32} to 10^{33} ergs down to the threshold for detection at about 10^{24} ergs. The energy release takes place on time scales of tens of seconds to minutes, corresponding to speeds of magnetic field annihilation as fast as 0.01 to 0.1 v_A. For the characteristic temperature and spatial scales of loops and arcades in the corona, the ratio of the resistive and Alfvén time ranges from 10^8 to 10^{14}. The characteristic time for the release of magnetic energy by reconnection in the Sweet-Parker model greatly exceeds the Alfvén time and is much longer than that inferred from observations.

The failure of the MHD model to explain the solar observations might be a consequence of the failure of the MHD equations to correctly describe dissipative phenomena in the highly conducting corona. The low values of resistivity lead to current layers with characteristic transverse scale lengths of $\sim L(\tau_A/\tau_r)^{1/2}$, which may be as small as the cyclotron radius of the ambient ions (~50 cm in the solar corona). The magnetohydrodynamic formulation of the dynamics is not valid at such scales and motivates the exploration of reconnection using kinetic models. In this low-collisionality regime the current density may also be sufficient to drive the electron conduction velocities above the ion thermal velocity, where strong excitation of plasma turbulence is possible. Scattering of electrons by the associated electric field fluctuations may greatly increase the effective resistivity and produce the "anomalous resistivity" that has been widely invoked in the literature. Whether such anomalous resistivity could also play a role in the production of such large numbers of energetic electrons (see Chapter 6) is not known. Exploration of these issues is ongoing.

Magnetic reconnection and the associated release of energy are believed to underlie other phenomena in the corona. For example, magnetic reconnection may be the ultimate source of heat in coronal holes (micro-flares), and so the origin of the solar wind, as well as the heat source for the x-ray-emitting corona (nano-flares), confined in the bipolar magnetic fields of both the ordinary and the ephemeral active regions. It is important to realize, however, that the form the energy release from reconnection takes is not limited to explosive solar-flare-type events. Indeed, the dominant process for coronal heating may be more gradually dissipative, or may arise from waves excited from outflows from reconnective events. A major scientific challenge is to understand the small-scale dynamics of the formation and internal structure of the current sheets arising from the essentially three-dimensional magnetic interactions that drive such reconnective energy release. Both theoretical studies and observations pushed to the highest resolution that technology can provide are essential for addressing the issues.

Outstanding Questions About Reconnection in the Solar Corona

- What controls the onset of solar flares and coronal mass ejections (see Chapter 5)?
- What is the lower limit on explosive flares in the corona?
- What physical mechanisms are responsible for particle energization during solar flares (see Chapter 6)?
- What are the dominant processes responsible for heating the corona?

MAGNETIC RECONNECTION IN EARTH'S MAGNETOSPHERE

Magnetic reconnection occurs in two general regions of geospace: at the magnetopause, the boundary that separates Earth's magnetosphere from the solar wind (or, more precisely, from the shocked and heated solar wind plasma of the magnetosheath); and in the magnetotail, the extended magnetic structure on Earth's nightside that stretches far beyond the Moon's orbit (see Figure 2.4). Reconnection at the magnetopause "opens" the geomagnetic field through the merging of a portion of the terrestrial field with the magnetic field entrained in the solar wind flow—the interplanetary magnetic field (IMF)—resulting in field lines that have one foot on Earth and the other on the Sun or in interplanetary space. Nightside reconnection closes Earth's field again through the merging of these open field lines. Magnetic reconnection is the principal mechanism by which energy, mass, and momentum are transferred from the solar wind to the magnetosphere and by which magnetic energy stored in the magnetotail is released in explosive events known as magnetospheric substorms. It thus plays a prominent role in the dynamics of Earth's magnetosphere.

Magnetopause Reconnection

Reconnection with the IMF is generally always occurring to some extent at the magnetopause, so that the magnetosphere is rarely completely closed. Where reconnection occurs on the magnetopause and how efficiently it effects the transfer of energy, mass, and momentum to the magnetosphere depend on the

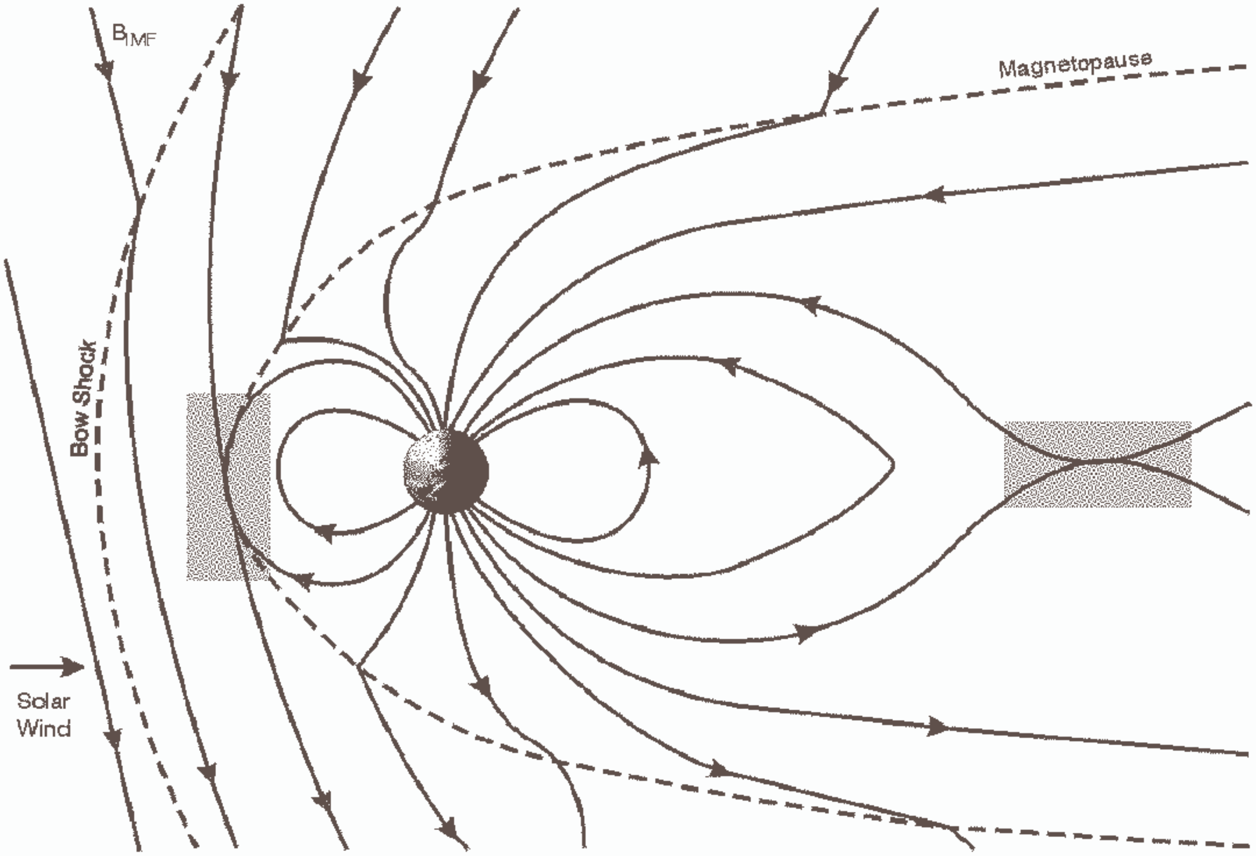


FIGURE 2.4 Southward-oriented interplanetary magnetic field (IMF) lines (blue) merge or reconnect with Earth's closed field lines (green) at the subsolar point. The merged or "open" flux tubes (red), with one end in Earth's ionosphere and the other end in the solar wind, are carried downstream by the solar wind flow and eventually reconnect in the distant tail. Merging results from the breaking of the frozen-in-flux condition, which occurs at an x-line in the diffusion or dissipation region (grey boxes). Merging of closed field lines in the near-Earth region of the magnetotail (not shown here) is associated with the onset of the substorm expansion phase. Reconnection at the dayside magnetopause is the primary mechanism for the transfer of mass, momentum, and energy from the solar wind to the magnetosphere, which occurs most efficiently when the IMF is oriented southward. In the tail, merging plays a role in the dissipation of energy stored in the magnetotail lobes as a result of dayside reconnection. (The drawing is not to scale.)

orientation of the interplanetary field relative to the geomagnetic field. In the simplest picture of magnetopause reconnection, the IMF is strongly southward—that is, it has an out-of-the-ecliptic component that is anti-parallel to Earth's northward-directed field at the subsolar magnetopause—and merges with the geomagnetic field across an extended portion of the dayside magnetopause, producing open field lines that are swept back into the magnetotail by the solar wind flow as shown in Figure 2.4.

Spacecraft and ground-based observations indicate that the onset of magnetopause reconnection is closely associated with the formation of large-scale, organized plasma flows in the ionosphere. These flows represent the motion of the ionospheric footpoints of the magnetic field lines that are undergoing reconnection at the magnetopause and later in the magnetotail and indicate how magnetic flux is replenished at the dayside magnetopause. As open field lines formed during magnetopause reconnection are transported into the tail by the solar wind flow, their footpoints move across the polar cap, from the dayside to the nightside. Subsequent reconnection in the magnetotail causes these open field lines to again become closed. They then contract toward Earth and flow around the flanks toward the dayside, where they resupply the dayside with magnetic flux. These flow signatures are now well reproduced by global magnetohydrodynamic models of the magnetosphere.

Satellite crossings of the magnetopause have yielded a wealth of data that document many of the phenomena predicted to result from reconnection, thus confirming both the occurrence of reconnection and its important role in the dynamics of the magnetosphere. These observations include direct measurement of plasma outflows from the reconnection site and magnetic field measurements that have verified predictions regarding the magnitudes and directions of these flows. Pairs of satellites flying on either side of the magnetic x-line have measured the expected oppositely directed outflows. High-time-resolution data from recent satellite observations has permitted the first exploration of the small-scale kinetic structure that has been predicted by theory to facilitate reconnection in the nearly collisionless environment of Earth's magnetosphere. Finally, direct measurement of the mixture of hot plasma from Earth's magnetosphere and the colder but denser plasma from the shocked solar wind on a single magnetic field line confirms that open field lines form as a result of magnetopause reconnection.

Because the IMF is generally not oriented directly southward but has a finite east-west component, the notion of oppositely directed field lines reconnecting at the subsolar magnetopause is an oversimplification. The location of magnetic reconnection at the magnetopause varies, depending on the direction of the IMF. Identifying the location of reconnection and understanding the physical processes that determine where reconnection takes place on the magnetopause continue to spark intense discussion in the scientific literature. The central issue is whether reconnection occurs primarily where Earth's field and the IMF are anti-parallel or whether reconnection can occur in regions where the magnetic field rotates through a finite angle (less than 180 degrees) across the magnetopause. In the latter case, called component reconnection, the magnetic field can be separated into a component that undergoes reconnection (within a defined plane) and a passive component perpendicular to the plane of reconnection. Component reconnection is generically the most common form of reconnection in the solar corona, astrophysical, and laboratory plasmas. For a given orientation of the IMF, there are always locations on the magnetopause where the IMF and magnetospheric magnetic field are oppositely directed. In the case of a nearly east-west-directed IMF, for example, the locations of anti-parallel fields are on the flanks of the magnetopause. There is some evidence from analyses of spacecraft observations that magnetic reconnection is favored in locations where the magnetosheath and magnetospheric magnetic fields are nearly anti-parallel and tracks these regions as the IMF direction changes in time.

Magnetotail Reconnection

The addition of magnetic flux in the tail lobes as a result of reconnection at the dayside magnetopause compresses and thins the magnetotail, producing an extended magnetotail current sheet. Threading through this current sheet is a small component of the magnetic field. This normal magnetic field inhibits magnetic reconnection (which would usually be expected to develop rapidly in a simple one-dimensional model) and therefore facilitates the buildup of flux and energy in the tail lobes. The pileup of magnetic flux in the tail can continue for long periods of time (up to several hours) during extended periods of magnetopause reconnection. Eventually, the formation of a magnetic x-line in the near-Earth region of the magnetotail leads to the onset of reconnection in the tail. Reconnection in this region either can be spatially and temporally localized or can organize into a large-scale event (a substorm). In the latter case reconnection proceeds until a significant fraction of the open flux that has built up in the tail reconnects. Field lines on the earthward side of this near-Earth x-line again become closed. At the same time, the field lines on the tailward side form disconnected magnetic flux tubes (see discussion in Chapter 5) that convect in an anti-sunward direction down the tail, disposing of the excess magnetic flux. Associated with this anti-sunward convection is the transport of plasma away from Earth, effectively reducing the plasma content of the closed field portion of the magnetotail. Through this process, the magnetosphere completes the cycle of loading and unloading of magnetic flux in the lobes initiated by reconnection at the magnetopause.

The development of a large-scale reconnection event that releases a substantial amount of the magnetic flux built up in the magnetotail is referred to as a substorm. The trigger mechanism for substorms remains uncertain and a number of competing theories have been proposed. Irrespective of the mechanism for the onset (see Chapter 5), satellite observations support the formation of a magnetic x-line (or lines) at distances of around 20 to 30 Earth radii—and in some cases as close as 15 Earth radii—anti-sunward from Earth. The transport and pileup of magnetic flux earthward of the x-line lead to a reconfiguration of the tail magnetic field and therefore the release of the magnetic stress associated with the stretching of the field lines by the solar wind flow. Anti-sunward of the reconnection region, the one or more reconnection sites combine to create plasmoids, large-scale traveling plasma structures entrained in magnetic flux ropes.

Magnetic reconnection in the tail current sheet does not necessarily develop as a long-term, large-scale phenomenon that releases a significant fraction of the energy stored in the lobes. Rather reconnection can be bursty and spatially localized. The flow signatures of such localized reconnection events, as measured by satellites, have been termed "bursty bulk flows." The earthward-directed flows from these bursts of reconnection transport flux toward Earth. While each individual event is small, the net transport from many such events is a major source of flux transport in the magnetotail. The physical processes that lead to such localized reconnection events and that limit their amplitude are not well understood.

Of all of the planetary bodies, it is only at Earth that reconnection has been extensively studied. The preceding discussion has therefore focused on the terrestrial case because of the relative abundance of data. However, it should at least be noted in conclusion that the Mariner 10 probe to Mercury and the Galileo probe to Jupiter have provided evidence for the occurrence of reconnection at those planets as well. Mariner 10 data have been interpreted as evidence for the occurrence of substorms in Mercury's tiny magnetosphere, while Galileo has observed the signature of what is likely to be the reconnection of stretched field lines in Jupiter's magnetodisk.

Outstanding Questions About Reconnection in Earth's Magnetosphere

- What are the relative roles of component reconnection and anti-parallel reconnection at the magnetopause? What determines the location of magnetic reconnection?
- Do coherent kinetic effects or turbulent scattering break the frozen-in condition during reconnection in the collisionless magnetosphere?
- What controls the onset of substorms (large-scale reconnection events in the magnetotail)?
- What controls the rate of magnetic reconnection and its spatial scale? Is reconnection steady or bursty?

THE ROLE OF LABORATORY EXPERIMENTS

The development of both ground- and space-based techniques for studying the dynamo and reconnection in the local cosmos combined with the development of theoretical/computational models has led to unprecedented progress in the understanding of both of these fundamentally important processes. Nevertheless, understanding naturally occurring dynamos and reconnection processes is complicated because of, for example, the complexity of the geometries, the inhomogeneity of important parameters, and the multiplicity of spatial scales involved. In recent years dedicated laboratory experiments have begun to play an increasingly important role in unraveling some of the important issues on these topics. Laboratory experiments have the advantage over naturally occurring phenomena that parameters can be varied to test ideas about the scaling of phenomena. Laboratory experiments on magnetic reconnection in particular have been constructed at national laboratories and university sites in both the United States and abroad. These experiments are now able to explore magnetic reconnection in both the collisional and collisionless regimes, test ideas about the scaling of the size of the dissipation region with parameters, explore the differences between reconnection with and without a guide field, and study the development of turbulence and its impact on the rate of reconnection. Theoretical modeling has in particular served to catalyze the interaction between laboratory experiments and satellite and other observations by providing testable ideas about the dominant processes that control reconnection. Several laboratory liquid metal dynamo experiments have also been constructed. Flows generated by propellers have been shown to reduce the rate of decay of seed magnetic fields, providing hope that the construction of larger-scale experiments (with larger Reynolds number) will demonstrate self-generation. An experiment that self-generates a seed magnetic field as a result of externally supplied flows would provide a wealth of data for benchmarking theoretical models.

CONCLUDING REMARKS

The generation of magnetic fields and their subsequent conversion into plasma kinetic energy have abundant examples throughout the universe. Thus, the creation and annihilation of magnetic fields take place over an enormous range of plasma densities and temperatures. However, in most cases similar physical processes are expected to control the essential dynamics. Solar physics and space physics are in a unique position to advance our understanding of these phenomena because of the accessibility of the Sun and the heliosphere to experimentation.

In the case of the Sun, high-resolution optical measurements can be used to investigate the small-scale fibril structure of the magnetic field and the role of magnetic reconnection in the development of flares and coronal mass ejections. Throughout the heliosphere, and especially at the planets, direct measurements of magnetic and electric fields, plasmas, and energetic particles can be used to test theories of the creation and annihilation of magnetic fields. Thus, the heliosphere is at once the setting for direct investigation of specific processes important to solar system plasmas and a laboratory for the investigation of magnetic-field phenomena important to the broader astrophysical plasma physics program.

NOTES

1. The strength of Earth's magnetic field is ~0.3 Gauss (30,000 nT) at the equator and twice that at the poles.
2. J. Bloxham and P.H. Roberts, The geomagnetic main field and the geodynamo, *Reviews of Geophysics*, Supplement, 428-432, 1991; P.H. Roberts and G.A. Glatzmeier, Geodynamo theory and simulations, *Reviews of Modern Physics* 72, 1081, 2000.
3. E.A. Spiegel and J.-P. Zahn, The solar tachocline, *Astronomy and Astrophysics* 265, 106-114, 1992.

4. N.O. Weiss, Solar and stellar dynamos, pp. 59-95 in *Lectures on Solar and Planetary Dynamos*, M.R.E. Proctor and A.D. Gilbert, eds., Cambridge University Press, Cambridge, United Kingdom, 1994.

5. G.A. Glatzmeier and P.H. Roberts, A three-dimensional convective dynamo solution with rotating and finitely conducting inner core and mantle, *Physics of the Earth and Planetary Interiors* 91, 63-75, 1995; and W. Kuang and J. Bloxham, An Earth-like numerical dynamo model, *Nature* 389, 371-374, 1997.

6. The origins of reconnection theory are reviewed by E. Priest and T. Forbes in the introduction to their book, *Magnetic Reconnection: MHD Theory and Applications*, pp. 6-10, Cambridge University Press, Cambridge, United Kingdom, 2000.

7. J. Birn, J.F. Drake, M.A. Shay, B.N. Rogers, R.E. Denton, M. Hesse, M. Kuznetsova, Z.W. Ma, A. Bhattacharjee, A. Otto, and P.L. Pritchett, Geospace Environmental Modeling (GEM) magnetic reconnection challenge, *Journal of Geophysical Research* 106, 3715, 2001.

3

Formation of Structures and Transients

Solar system and astrophysical plasmas exhibit dynamic behavior whenever different plasma regimes interact with one another. Cosmic plasmas from diverse sources generally populate uniquely defined volumes of space that border against similar volumes populated by plasmas from other sources. Such cosmic plasmas are generally magnetized, and so different plasma regimes resist interpenetration. Examples of plasma regime interactions that occur in the solar system are (1) interactions between coronal mass ejections and the background solar wind, (2) interactions between the solar wind and the ionospheres and magnetospheres of planets, (3) interactions between the corotating, subsonic planetary magnetospheric plasmas and the small magnetospheres surrounding planetary satellites (e.g., the interaction of Ganymede with the magnetospheric plasma of Jupiter), and (4) interactions between the solar wind and the local interstellar medium at the outer reaches of the solar system. As is discussed here, such examples have clear relevance and application to many extrasolar astrophysical plasma systems.

When the relative velocity between plasma regimes is supersonic, the first interaction boundary is a *collisionless shock*. Whether or not the interaction is supersonic or subsonic, *current sheets* generally separate the different plasma regimes. The stresses imposed by the interactions also engender the formation of current sheets within the plasma regimes, leading to a *cellular structure*. Such internal and boundary current sheets imply the existence of high concentrations of energy density and shear stress, and the system responds to dissipate and redistribute the energy and the stress, for example, through magnetic reconnection. These redistributions engender structuring within current sheets, with features such as *magnetic flux ropes*. The energy redistribution and structuring are accomplished in part by a fundamental property of plasmas, *cross-scale coupling*. This coupling connects small scales, where particle kinetic effects are important, to large scales, creating features like boundary layers at the walls of cells. Dissipation of kinetic-scale structures can yield charged particle heating and particle acceleration. A special case of cross-scale coupling is *hydromagnetic turbulence*, which results in the generation of numerous and hierarchically ordered spatial and temporal scales and in the transfer of energy to smaller and smaller spatial scales. The sections that follow discuss each of these structures in turn and identify some of the outstanding questions associated with them. The chapter concludes with a discussion of their universality.

COLLISIONLESS SHOCKS

In an ordinary gas, a shock wave is created in front of an object that is moving at a supersonic velocity (i.e., at a speed with a Mach number greater than 1) relative to the gas. This shock converts flow energy into thermal energy (heat) through a dissipation mechanism. In a gas, this dissipation mechanism is the collisions between gas particles. By analogy, a shock wave should be produced in front of an object, such as a planet, that is immersed in a supersonically flowing plasma such as the solar wind. Forty years ago, the analogue of a hydrodynamic shock wave in a plasma was a hotly debated topic. At the heart of this debate was the dissipation mechanism in a shock where the mean free path of a particle (i.e., the distance over which a particle will probably suffer a collision with another particle) was much larger than the scale length over which any dissipation must take place. For example, in the supersonic solar wind flow past Earth, the mean free path of a particle is about 1 astronomical unit (AU; 1.5×10^8 km), while the dissipation scale length is predicted and observed to be of the order of 100 km. The discovery of Earth's bow shock (Figure 3.1) in the early 1960s settled the debate as to the existence of such "collisionless" shocks and raised the issue of dissipation mechanisms to the forefront.

In the intervening 40 years, shocks have been identified in the interplanetary medium, at other planets, at comets, and (indirectly) at the boundary between the heliosphere and the interstellar medium. Moreover, a strong interplay between analytic theory, computer simulations, and in situ observations has led to a remarkable understanding of the dissipation mechanisms in collisionless bow shocks. Significant advances in our knowledge of Earth's bow shock and in our theoretical understanding of collisionless shocks were achieved during the 1980s in particular, with the demonstration of the importance of the magnetic field and plasma kinetic effects in shock dissipation.[1]

An important development during this period was the discovery of a population of reflected ions at quasi-perpendicular shocks—that is, shocks where the angle between the solar wind magnetic field and the shock normal is greater than 45 degrees (cf. Figure 3.1). With the improved observations, the shock structure on ion gyroscale lengths was resolved; and within a gyroradius of quasi-perpendicular shocks, a portion of the solar wind ion distribution was observed to reflect off the shock, gain energy in the upstream region, return to the shock, and enter the downstream region. These reflected ions were found to play an important role in the dissipation process at supercritical shocks—that is, shocks, like planetary bow shocks, where a certain critical Mach number is exceeded. The identification of the reflected ion population resulted from a significant improvement in the time resolution of in situ plasma observations at the bow shock, while state-of-the-art computer simulations provided an understanding of the process by which ions are reflected at supercritical shocks. Hybrid simulations, where the ions are treated as particles and the electrons are treated as a fluid, showed that the reflection process requires both electric and magnetic fields.

Since the early 1980s, the structure of more complicated shocks has been investigated. In particular, dissipation in quasi-parallel shocks, for which the angle between the solar wind magnetic field and the shock normal is less than 45 degrees, was investigated at Earth's bow shock in the late 1980s. For quasi-parallel shocks, upstream waves generated by ions propagating away from the shock play an important role. Because of these waves, the quasi-parallel shock undergoes periodic overturning and reforming. Observations showed that during the reformation process, the shock dissipation is similar to that observed at quasi-perpendicular shocks. The enormous success of shock research in the 1980s has provided a base of understanding of some more exotic shocks elsewhere in the solar system and in the universe (Figure 3.2).

In addition to heating the plasma, shocks accelerate a fraction of the particle population.[2] The presence of a population of energized particles at a shock can, in turn, appreciably influence the shock structure. Specifically, the energized particles provide at least some of the dissipation that is required for

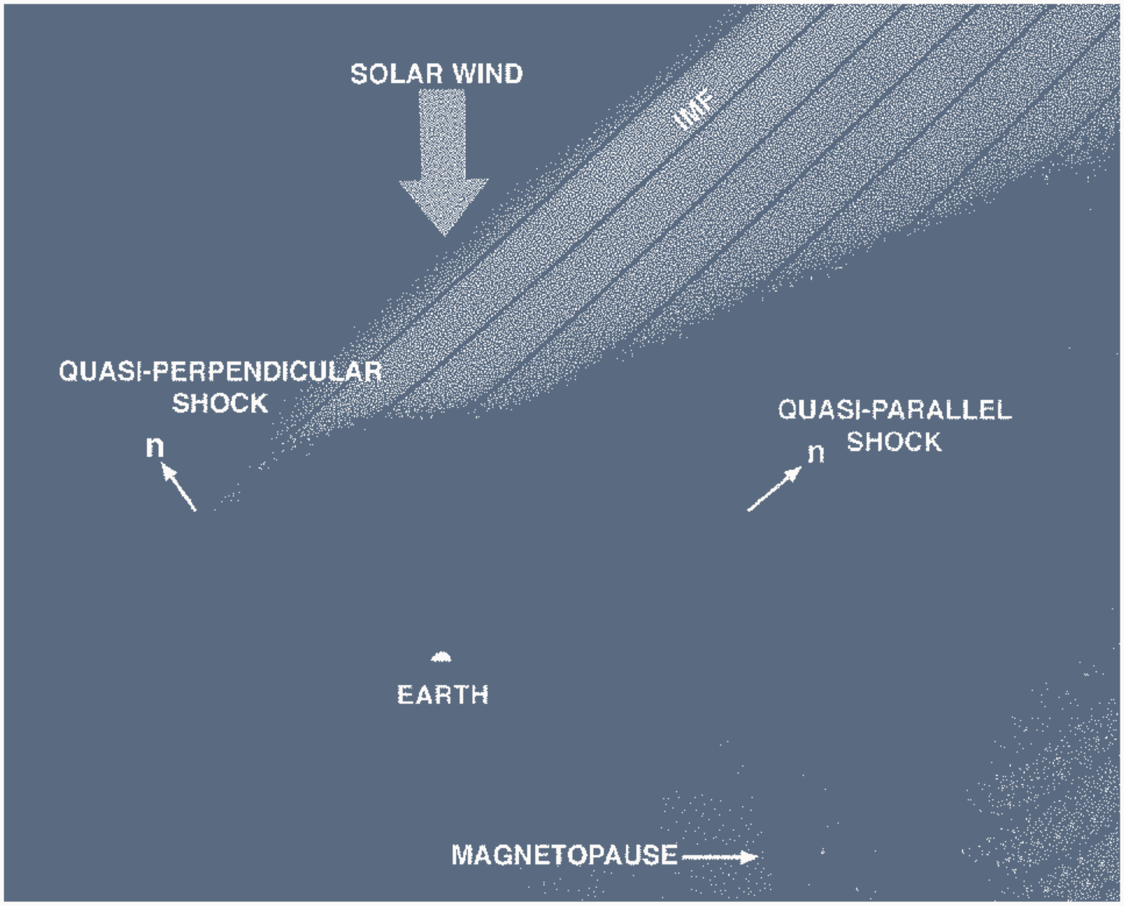


FIGURE 3.1 Schematic showing Earth's quasi-perpendicular and quasi-parallel bow shock. Shock normal is indicated by the arrow labeled "n." Adapted, with permission, from B.T. Tsurutani and P. Rodriguez, Upstream waves and particles: An overview of ISEE results, *Journal of Geophysical Research* 86(A6), 4319-4324, 1981. Copyright 1981, American Geophysical Union.

the shock to satisfy the Rankine-Hugoniot conditions (the mathematical conditions that describe the changes in plasma and magnetic field parameters across discontinuities), and thus help establish the structure of the shocks. Also, the energized particles can become major participants in the balance of energy and momentum across the shocks. Given the potentially large gyroradii of such particles, the structure of the shocks can be further modified and thickened.

Studies of Earth's bow shock have resulted in an excellent base of understanding for shock dissipation and shock acceleration in collisionless plasmas. However, the range of parameters represented by planetary bow shocks is limited. A critical uncertainty concerns the influence of particle acceleration and other dissipation processes on the structure of different kind of shocks, including (1) shocks associated with comets, where pickup and other neutral gas interactions play a role; (2) the solar wind termination and heliospheric boundary shocks, where strong particle acceleration may play a role; and (3) very strong astrophysical shocks associated with supernova remnants.

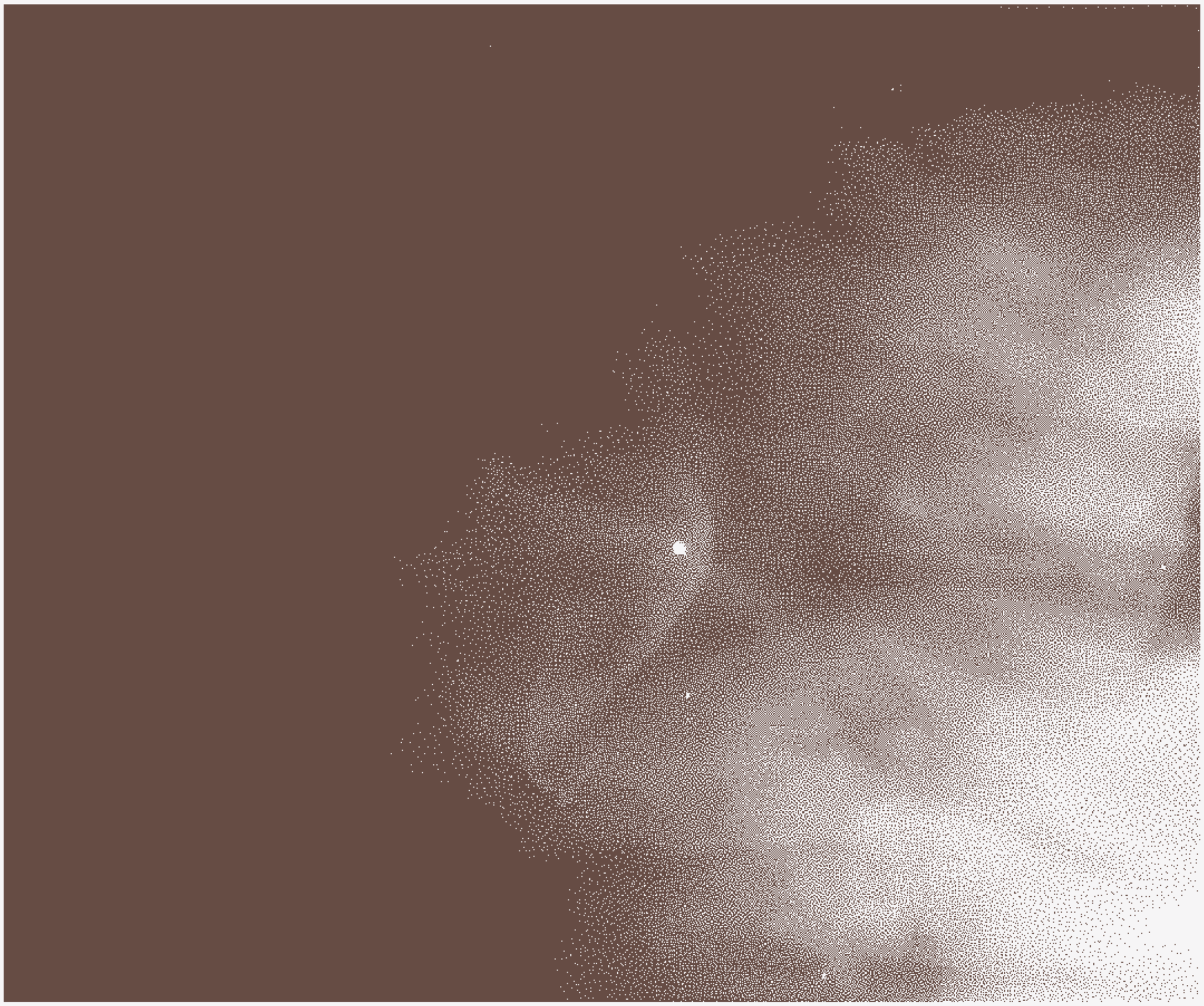


FIGURE 3.2 Bow shock upstream of the young star LL Ori, which is located some 1500 light-years from Earth in the star-forming region of the Orion Nebula. Courtesy of C.R. O'Dell (Vanderbilt University) and the Hubble Heritage Team (NASA/STScI/AURA).

Outstanding Questions About Collisionless Shocks

- How do strong particle acceleration and associated energetic particles modify shock structure?
- How do pickup and other neutral gas interactions modify shock structure and particle acceleration at comets and heliospheric boundary shocks?
- How do researchers extrapolate knowledge of shock structure derived from studies of solar system shocks to the strong shocks that prevail in astrophysical systems?

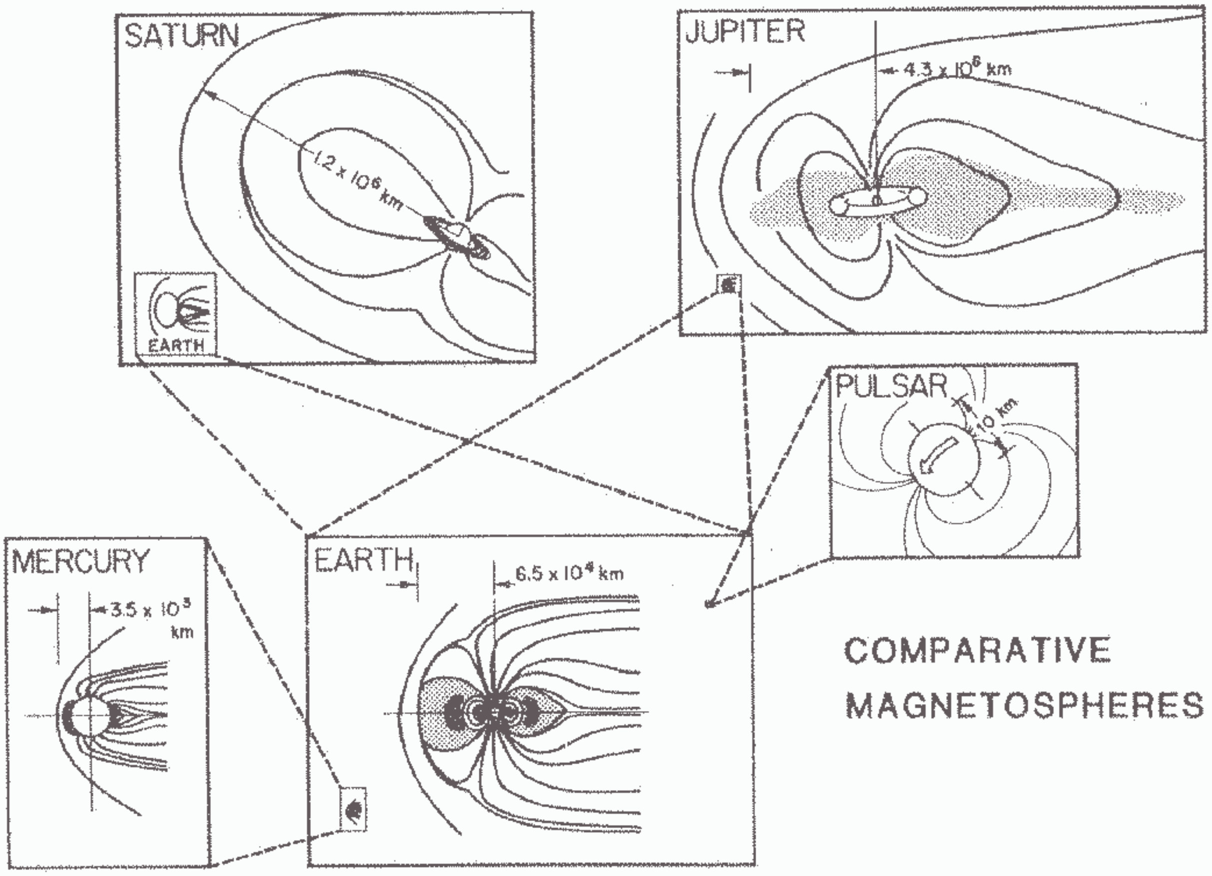


FIGURE 3.3 Magnetospheres epitomize the cellular organization of astrophysical plasmas.

CELLULAR STRUCTURES AND CURRENT SHEETS

Cellular Structures

Magnetized plasmas in space tend to form cells enclosed by current sheets. "We find that space occupied by plasmas tends to be divided in 'compartments,' often separated by thin current sheets. On the interstellar and intergalactic scale, space may have a cellular structure consisting of many separate regions, containing plasmas of different magnetization, density, temperature, and perhaps also different kinds of matter."[3] Figure 3.3 shows a hierarchy of magnetic cells in the form of various magnetospheres.

Cells form when magnetized plasmas from spatially separated origins, pursuing their natural, expansionist tendencies, collide, or compete for the same space. The surface along which the competing plasmas meet, which ideally is a current sheet, can be thought of as the cell wall. If the competition produces a stationary (or quasi-stationary) standoff with one side enveloping the other, the notion of a cell is especially apt.

Planetary Magnetospheres

In the solar system, planetary magnetospheres epitomize cellular structures. When the solar wind encounters a strongly magnetized planet that is surrounded by a plasma, the boundary that separates the solar wind and the localized plasma is called a magnetopause. Magnetopauses exist at Earth, Jupiter, Mercury, Saturn, Uranus, and Neptune (Figure 3.4) and should exist at any magnetized body immersed in a stellar wind. Inside the magnetopause is a vast volume that is dominated by the plasma and magnetic field associated with the planet, while outside the magnetopause the solar wind's plasma and magnetic field dominate. Similarly, a magnetopause can separate planetary magnetospheric plasmas from those confined to the vicinity of a planetary satellite, as in the case of Ganymede in Figure 3.4.

For Earth's magnetosphere, the problem of the gross structure of the magnetopause was first solved in the early 1960s in a restricted form known as the Chapman-Ferraro solution. In this restricted form, the size and shape of the magnetopause are determined by the ram pressure of the solar wind and the strength of the planetary magnetic dipole and its orientation relative to the solar wind. Chapman-Ferraro scaling for size is well satisfied for Mercury, Saturn, Uranus, and Neptune. Jupiter deviates markedly from Chapman-Ferraro scaling. At this massive planet, internal particle pressure plays a major role and therefore violates the Chapman-Ferraro assumption that the magnetosphere is a vacuum. Chapman-Ferraro scaling does not apply to Venus and Mars because they lack significant internal magnetic fields. Although Venus, Mars, and comets do not satisfy this Chapman-Ferraro scaling, they do generate a cellular structure with a leading shock because they are immersed in the supersonic solar wind.

The Heliosphere

The entire solar system resides within a grand plasma cell called the heliosphere, which constitutes a bubble of plasma within the interstellar medium (Figure 3.5). The detailed interaction between the local interstellar medium (LISM) and the solar wind is not well understood, in part because many of the pertinent physical parameters of the LISM are poorly constrained. As the solar wind expands into the LISM, the complex boundaries of the bubble shield us from the interstellar plasma and magnetic fields and most of the cosmic rays and dust that compose the LISM. Indirect evidence and models suggest the following heliospheric boundaries. The supersonic solar wind eventually decelerates to subsonic speeds via a shock—the solar wind termination shock. This structure, to be encountered by the Voyager probes, our most distant spacecraft, is predicted to lie some 100 AU from the Sun. Beyond the termination shock, a boundary layer, the heliopause, exists were the shock-heated solar wind is cooled and diverted as it encounters the LISM flow. Because of uncertainty with respect to the role of the intergalactic magnetic field and very energetic particles in the LISM, it is currently not known whether the LISM flow is supersonic or subsonic. If supersonic, then, like the solar wind deceleration at a planetary shock, the LISM also decelerates via a shock, and the system is referred to as a two-shock model (i.e., bow shock plus termination shock). By contrast, a subsonic LISM does not require a bow shock, and the system is described as a one-shock model.[4]

Very detailed and sophisticated models that include the self-consistent coupling of plasma and neutral atoms (both interstellar and heliospheric) have been developed over the past several years, and these models are beginning to be used to investigate the stellar-wind-dominated regions (asterospheres) surrounding stars other than the Sun. In this respect, the very detailed plasma physics of the solar wind/LISM interaction has opened up a new field of astrophysics, in which recent observations of Lyman-α absorption profiles toward nearby stars have been interpreted in terms of the interaction of stellar winds with a partially ionized interstellar medium.[5]

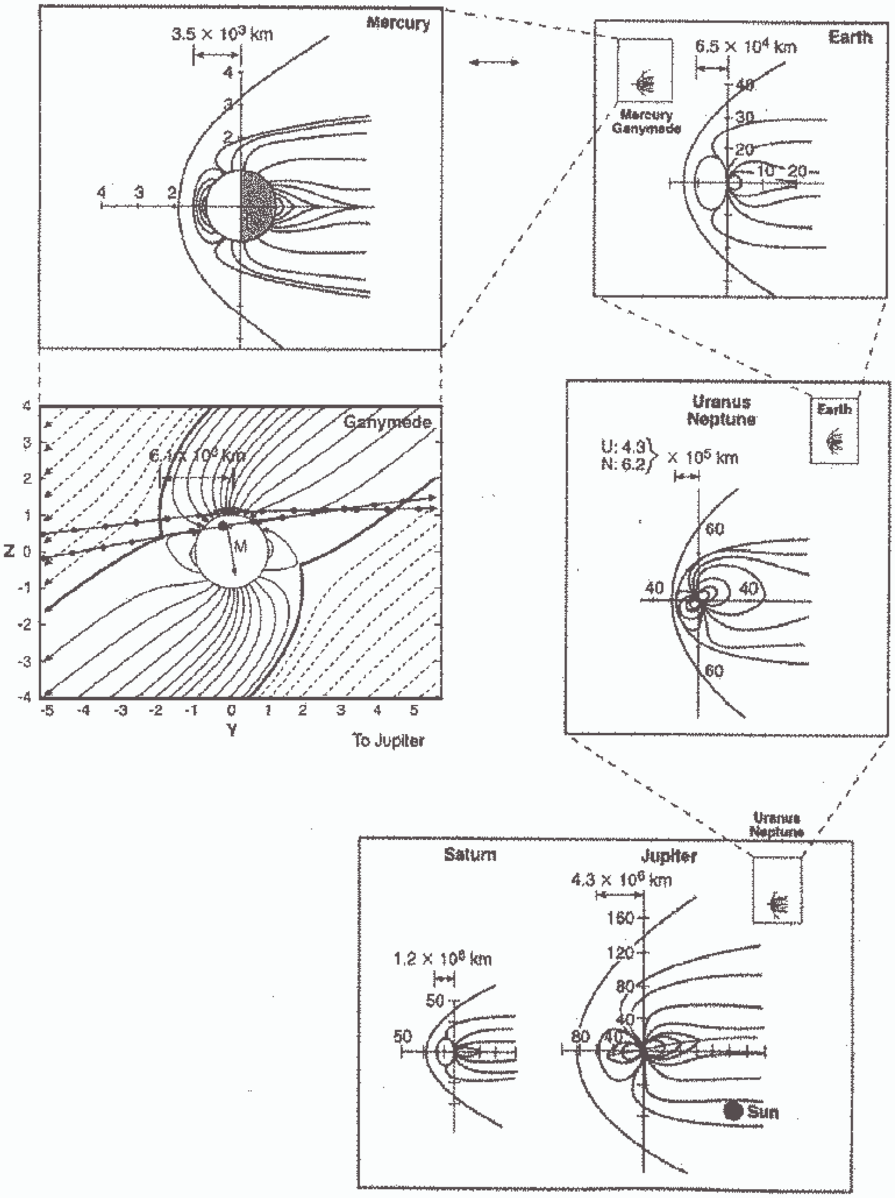


FIGURE 3.4 Schematic diagram showing the magnetospheres in the solar system that result from the interaction of the solar wind with the intrinsic magnetic field of the planets, and, for the case of Ganymede, from the interaction between Jupiter's magnetospheric plasmas and those confined to the vicinity of Ganymede. In the Ganymede schematic, the thick line separating dashed (magnetic field lines unconnected to Ganymede) and solid (Ganymede-connected) field lines constitutes the magnetopause analogue. The radii (in kilometers) of the planets, moon, and Sun represented are $R_M = 2436$, $R_G = 2631$, $R_E = 6378$, $R_N = 24874$, $R_U = 26150$, $R_S = 60272$, $R_J = 71434$, $R_{sun} = 695{,}990$. Note (bottom panel) the size of the Sun relative to the jovian magnetosphere, which is the largest object in the solar system. Courtesy of D.J. Williams (Applied Physics Laboratory). The panel showing Ganymede's magnetosphere is adapted, with permission, from Figure 3 of M.G. Kivelson et al., The magnetic field and magnetosphere of Ganymede, *Geophysical Research Letters* 24(17), 2155-2158, 1997. Copyright 1997, American Geophysical Union.

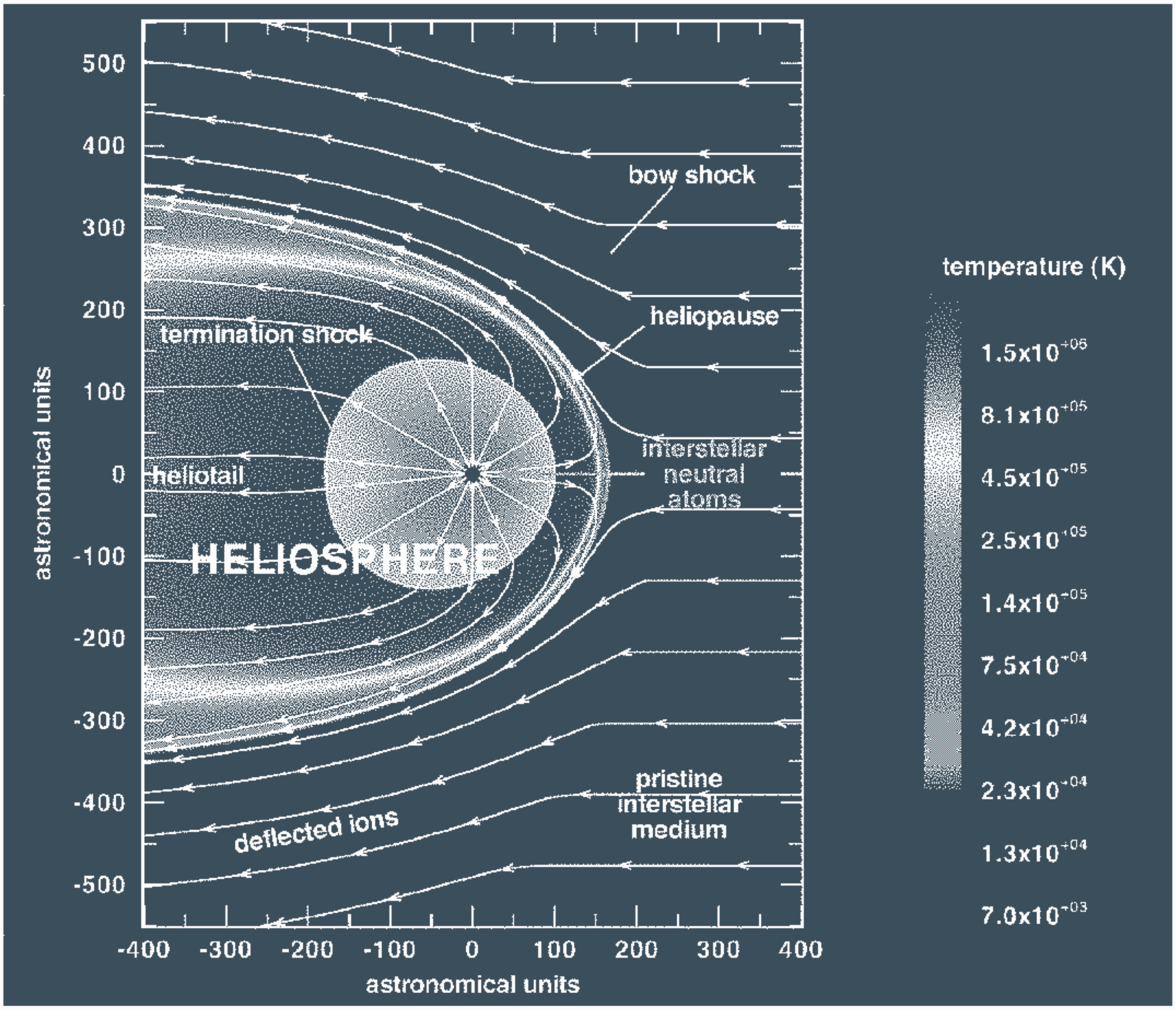


FIGURE 3.5 Simulation of the interaction of the heliosphere with the local interstellar medium (LISM). Whether a bow shock forms upstream of the nose of the heliosphere depends on whether the velocity of the LISM relative to the heliosphere is supersonic or subsonic, which is not known. The magnetized interstellar plasma is excluded from the heliospheric cavity and flows around it. However, interstellar neutral atoms can pass freely through the heliopause. Once in the heliosphere, some of them are ionized by charge exchange with solar wind protons, picked up by the solar wind, and carried back toward the termination shock. At the shock some of the pickup ions are accelerated to extremely high energies and return to the inner heliosphere as anomalous cosmic rays. Image courtesy of V. Florinski (University of California, Riverside). Reprinted, with permission, from V. Florinski et al., Galactic cosmic ray transport in the global heliosphere, *Journal of Geophysical Research* 108(A6), 1228, doi:10.1029/2002JA009695. Copyright 2003, American Geophysical Union.

Current Sheets

The controlling agents in the formation and maintenance of the cellular structure of solar system plasmas are current sheets. By definition, a current sheet carries a significant electrical current within a two-dimensional (sheet-like) structure. The term "current sheet" implies a structural lifetime much longer than characteristic plasma time scales (e.g., gyroperiod, plasma period), thus distinguishing it from plasma wave structures. And it implies a structure that moves only slowly with respect to the plasma, thus distinguishing it from shocks. The current in current sheets may be oriented either perpendicular to the magnetic field or parallel to it, and the distinction between perpendicular and parallel current sheets is not always a clean one.

Three distinct processes produce sheets of perpendicular current, in which the current density **J** is largely perpendicular to the magnetic field **B**. The first process involves the collision between two plasmas. The intervening current layer typically collapses to a sheet having a thickness close to the gyroradius of the dominant ion. Familiar examples include the dayside magnetopause of a planetary magnetosphere, the dayside ionopause of an unmagnetized planet (in analogy to the magnetopause at a magnetized planet), the boundary between solar-wind streams of different properties, and (presumably) the as-yet unexplored heliopause. A second process by which current sheets are produced is the stretching and dragging of magnetic flux tubes downstream by a flowing plasma. These flux tubes are anchored to a stationary obstacle in a background flowing plasma, forming two adjacent lobes of oppositely directed flux. The current sheet resides between the two adjacent lobes. This process occurs in planetary magnetotails, in cometary ion tails, and in the analogous induced magnetotails that appear downstream of an unmagnetized planet or satellite with a conducting atmosphere that is immersed in a flowing magnetized plasma (e.g., Venus, Mars, Titan). A third current-sheet production mechanism involves the inflation of a quasi-dipolar field by internally generated plasma stresses. An example of the agent of inflation can be the centrifugal acceleration of partially corotating plasma in a rapidly rotating magnetosphere like that of Jupiter and perhaps Saturn. A prominent example of the third mechanism is the heliospheric current sheet formed by the outflowing solar wind (Figure 3.6).

With the three basic kinds of dynamically created current sheets, the currents flow with a large component that is perpendicular to the magnetic field vectors on each side of the sheets. Sheets of parallel current (**J**||**B** to a good approximation) are also ubiquitous but arise from different causes and have different effects. An example is the current sheets associated with the northern and southern lights, the auroras.

Magnetopause current sheets arise naturally, for example, in all magnetohydrodynamic (MHD) models of planetary magnetospheres. Other current sheet structures are not so well understood. Jupiter's magnetodisk is one such structure. In general, it is not known which forcing terms (e.g., the centrifugal force) dominate in the fully developed magnetodisk and therefore how the magnetodisk current is generated. Pressure gradients, pressure anisotropies, field-aligned plasma acceleration, and beaming effects can all contribute, and the evolution of the current sheet can redistribute the importance of the various terms.

Across the domain of all current sheets, the greatest immediate challenge lies in the coupling that occurs between small and large scales, both in the temporal and spatial domains. Current sheets are the sites where the different interacting scales of a magnetized cosmic structure come together and influence each other.

Outstanding Questions About Current Sheets

- What factors determine the stability and instability of current sheets?
- How does Jupiter's unique magnetodisk form and how is it supported?

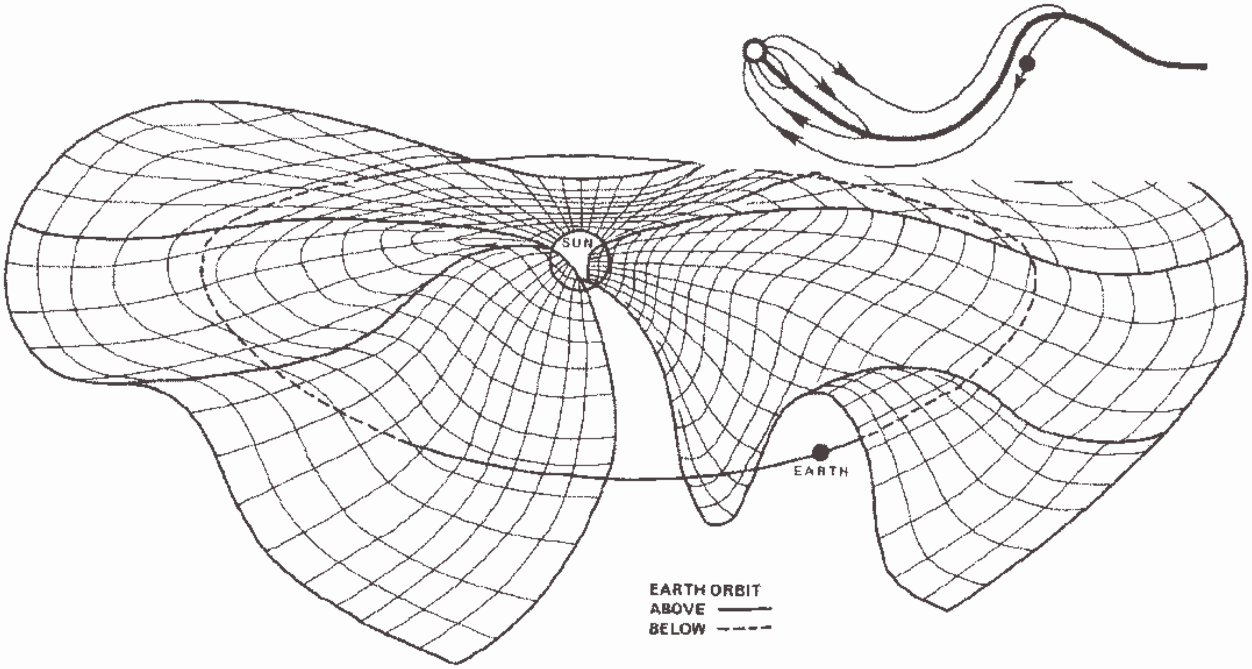


FIGURE 3.6 Schematic diagram of the three-dimensional structure of the current sheet that flows in an azimuthal direction around the Sun. The inset at the top of the figure shows the opposite polarities of the magnetic fields on the two sides of the current sheet. Courtesy of S.-I. Akasofu, Geophysical Institute, University of Alaska.

CURRENT SHEET STRUCTURING: BOUNDARY LAYERS AND FLUX ROPES

The separation of the plasma cells by current sheets and, where applicable, the accompanying shocks, is not perfect. Current sheets are often inherently dynamic, and they tend to form smaller-scale structures that are often dissipative. Dissipative processes allow energy and mass to be transported across the boundary. Processes that work toward the dissipation and destruction of current sheets are reconnection, tearing instabilities, impulsive penetration across boundaries by fast-flowing, high-density or high-pressure plasma blobs, the Kelvin-Helmholz instability, non-adiabatic particle acceleration, and the coupling to adjacent regions by field-aligned currents and the associated shear-stresses within the magnetic field. An important consequence of this structuring is the formation of boundary layers at current sheets. Examples of boundary layers formed by structuring are Earth's low-latitude boundary layer and high-latitude mantle. Such boundary layers often have sheared plasma flows, magnetic field-aligned particle streaming, and strong particle pressure gradients extending substantial distances from the current sheet proper.

Current sheet structuring can also cause some current sheets to periodically or sporadically disappear altogether. An example is the earthward diversion of the current sheet within Earth's magnetotail during the dynamical events called substorms. The current sheet dissipative processes listed above can generate specific classes of spatial structures. Chapter 2 discusses boundary layers generated at Earth's magnetopause and within Earth's magnetotail by magnetic reconnection. This section focuses on a specific class of current sheet structuring—flux ropes—because of its general applicability and importance for plasmas throughout the solar system, and presumably beyond.

A magnetic flux rope is a magnetic field configuration with the following characteristics: (1) locally tubular geometry, (2) helical magnetic field lines with zero twist on the axis and a pitch angle increasing from zero with increasing distance from the axis, and (3) maximum field strength on the axis. This qualitative geometrical definition of a magnetic flux rope does not specify the plasma characteristics, the global magnetic field topology, or the generation mechanism. Magnetic flux ropes can be found in the laboratory, in a variety of places in the solar system, and in other astrophysical settings. The geometry of the magnetic field in a flux rope implies the existence of a component of current along the magnetic field lines. These structures can be generated by magnetic reconnection and other plasma processes.

One of the earliest applications of flux ropes to solar physics was the Gold-Hoyle theory of solar flares, which proposed a model of flux ropes with constant twist.[6] Since then, the concept of flux ropes has grown enormously in importance for interpreting solar phenomena (Figure 3.7). Solar active regions themselves are now considered to be manifestations of large flux ropes. It is widely held that each bipolar active region is a single flux rope that has risen buoyantly through the convection zone from a dynamo layer.[7]

Flux ropes are observed in the solar wind as the magnetic field configuration within magnetic clouds. Magnetic clouds are defined as transient ejecta with (1) greater than average magnetic field strengths, (2) a smooth rotation in magnetic field direction, and (3) low proton temperatures and β (the ratio of plasma to magnetic pressure). They are a particularly well-organized type of coronal mass ejection. Globally, the magnetic clouds observed near 1 AU have the form of loops or ropes typically with both ends connected to the Sun (Figure 3.8).[8] Small magnetic flux ropes, with diameters 10 to 40 times smaller than those of magnetic clouds (i.e., 1-4 × 10^6 km), have recently been observed in the solar wind. These flux ropes differ from magnetic clouds in several respects, and they probably have an origin different from that of magnetic clouds.

More than 20 years ago flux ropes were discovered in the ionosphere of Venus.[9] These were observed as a series of large magnetic field strength enhancements, many with a ratio of maximum to minimum field strength of the order of 50. They are more numerous at lower altitudes, suggesting that either they form there and rise buoyantly or they form near the ionopause and are dragged toward Venus by the interplanetary magnetic field to which they are connected.

In the mid-1980s evidence was found for the existence of flux ropes in the distant geomagnetic tail, »100 R_E downstream from Earth.[10] The flux ropes were initially identified on the basis of a south-then-north tilt of the magnetic field, a strong core field, and a significant east-west component of the field in "plasmoids" moving at several hundred kilometers per second down the tail. An association with substorms was noted. Other flux rope structures have been observed more recently within the magnetotail. Several mechanisms for the formation of flux ropes in the magnetotail have been proposed,[11] but the favored hypothesis is magnetic reconnection at near-Earth and distant neutral lines.

Magnetic flux ropes have been discovered in many different locations in the solar system, and they undoubtedly occur in many other astrophysical systems. They may have a common geometrical form, but depending on boundary conditions and temporal evolution, they will differ to varying degrees from this form. Flux ropes and the boundary layers from which they can arise share a number of fundamental physical problems and questions.

Outstanding Questions About Boundary Layers and Flux Ropes

- How are mass and energy transported across collisionless boundary layers?
- What factors cause current sheets and boundary layers to form flux rope structures?

FIGURE 3.7 Computer simulation of a flux rope in the solar corona. The false color indicates the magnetic field strength in teslas. Image courtesy of I. Roussev (University of Michigan). Reprinted, with permission, from I.I. Roussev et al., A three-dimensional flux rope model for coronal mass ejections based on a loss of equilibrium, *Astrophysical Journal* 588, L45-L48, 2003. Copyright 2003, American Astronomical Society.

- Under what conditions are flux ropes stable and unstable?
- What is the relationship of flux ropes to reconnection?
- How do flux ropes evolve and what determines their sizes?
- How are flux ropes destroyed?

CROSS-SCALE COUPLING

Flux ropes are just one important class of structures that demonstrate the fundamental propensity of plasmas to couple strongly across multiple scales. This property of magnetized plasmas has important

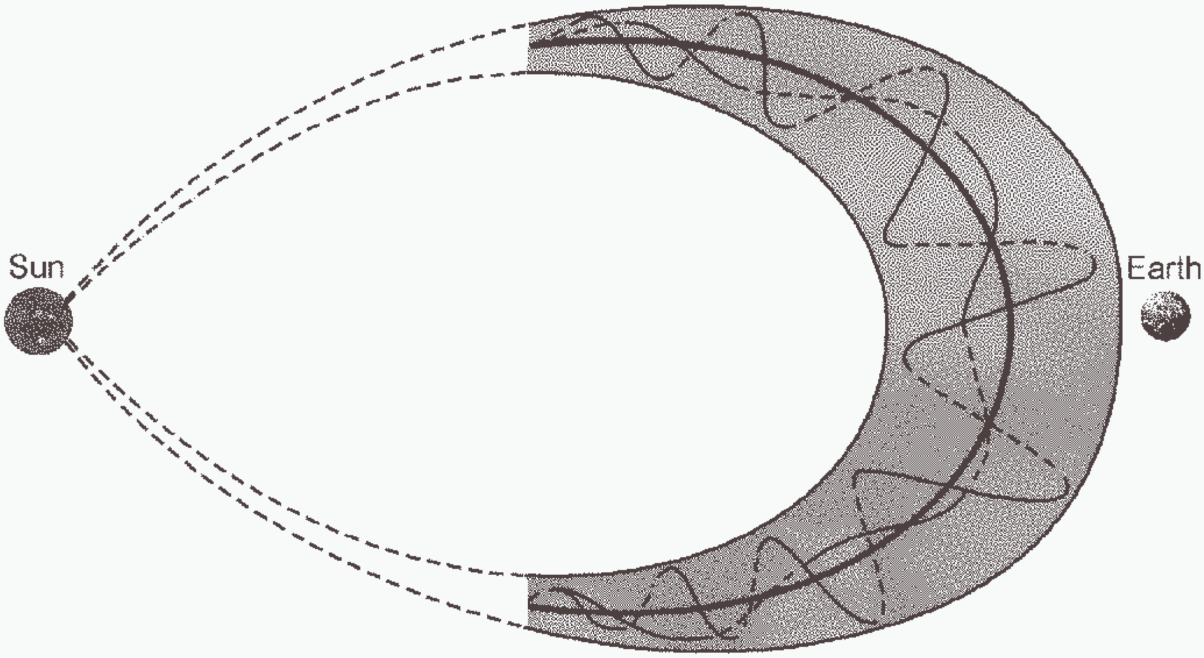


FIGURE 3.8 Magnetic flux rope in the form of a "magnetic cloud" from the Sun. Modified, with permission, from L. Burlaga, Global configuration of a magnetic cloud, pp. 373-377 in *Physics of Magnetic Flux Ropes*, Geophysical Monograph 58, C.T. Russell, E.R. Priest, and L.C. Lee, eds., American Geophysical Union, Washington D.C., 1990. Copyright 1990, American Geophysical Union.

implications for structures in plasmas and plasma dynamics. Cross-scale coupling in a plasma can be coherent or turbulent. In a coherent interaction, a specific wave mode generated by free energy in a plasma affects macroscopic plasma parameters by a microscopic, direct, resonant interaction (e.g., a wave-particle resonant interaction). Coherent coupling between microscopic and macroscopic scales presents a profound challenge to space plasma theory and modeling. It is often difficult to understand the microscopic plasma processes because they result from complicated boundary conditions imposed by the macroscopic system and because they are normally observed in a state of nonlinear saturation, which is difficult to treat theoretically. Furthermore, the microscopic process has a controlling influence on the large-scale situation, so that the system must be treated self-consistently. Examples of coherent coupling include magnetic reconnection, the generation of monochromatic Kelvin-Helmholtz waves on the magnetopause, and the generation of quiet auroral arcs from large-scale plasma flows.

At the opposite extreme is turbulent coupling, where many dynamical modes of the system are simultaneously stimulated and interact strongly. The most active states of the aurora, where no dominant spatial scale can be found, undoubtedly involve turbulent coupling. Although turbulence is in a saturated state, turbulent processes also evolve and must be treated over a large range of scales. Hydromagnetic turbulence, a special case of turbulent coupling in space plasmas, is discussed in some detail in this section as an example of the challenges to understanding this fundamental plasma process.

Hydromagnetic Turbulence

Probably the most well-studied example of cross-scale coupling in a plasma is hydromagnetic turbulence. Turbulence is a broadband, nonlinear dynamical interaction of fluctuating quantities (e.g., velocity) at multiple scales in a fluid, magnetofluid, or plasma. Larger scales tend to feed energy to smaller scales and this "cascade" process continues down to a dissipation scale, where heating occurs. As generally understood, the reason that energy transfers to small scales through the cascade process is simply that waveforms or structures steepen and stretch owing to one flow or current system shearing and compressing another.

Examples of turbulent dissipative processes are found in the small-scale magnetic interactions in the chromosphere and transition regions of the solar atmosphere and their possible coupling to the corona and solar wind. Furthermore, the solar wind itself has been observed to evolve toward a fully MHD turbulent state as it propagates toward the magnetosphere. In the magnetosphere, thin boundary layers such as the magnetopause are unstable to turbulent plasma processes. Nonlinear growth and saturation of these processes may lead to enhanced particle transport and heating. Turbulence apparently plays an important role in current disruptions and bursty bulk flows in Earth's magnetotail and in their correlation with large magnetic field fluctuations. Finally, turbulence plays an important role in dynamo processes (see Chapter 2).

Turbulence has been studied most completely in the solar wind.[12] The characteristic spectra of turbulence are observed over a few decades of scales in solar wind magnetic fields, velocities, densities, and temperatures, and these spectra evolve with distance from the Sun. Fundamental questions concerning turbulence remain even as it relates to the well-studied solar wind. These questions relate to turbulence spectra and their evolution, cascade rates, symmetry, and "Alfvénicity" correlations.

The most frequently cited characteristic of turbulence is a $k^{-5/3}$ spectrum of any quantity (e.g., velocity) as a function of the magnitude of the wave vector, k, in the medium. Kolmogoroff derived this spectrum in 1941 for the velocity field in an isotropic, high-Reynolds-number (low-viscosity) fluid.[13] He used dimensional arguments for the "inertial range" of wave numbers (k's) in which viscosity is unimportant compared to nonlinear terms. Viscosity set the dissipation scale, and the large, "energy-containing" scales decayed at a slow but predictable rate. For unknown reasons, the –5/3 spectrum is observed in the solar wind despite the fact that it is inhomogeneous, is anisotropic, and contains a magnetic field that could slow the interactions and flatten the spectrum.

The turbulent spectral level determines the steady-state cascade rate of energy from one scale to the next in the Kolmogoroff formalism. Energy conservation implies that the cascade rate is equal to the dissipation rate. These turbulence constraints can be used to compare predicted and observed heating of the solar wind. In general, predictions and observations agree reasonably well with respect to the evolving solar wind in the inner heliosphere. However, attempts to use phenomenological cascade rates to account for heating of the corona and acceleration of the solar wind suffer from uncertainties in the fluctuation levels. In particular, it is not clear that fluctuations can be generated at high enough levels to account for the observed coronal heating.

The magnetic field plays an important role in creating symmetry. Fluctuations with wave vectors along the mean magnetic field are much more effective in scattering particles than are those with wave vectors nearly transverse to the field. The interplanetary fluctuations may contain significant levels of the quasi-two-dimensional fluctuations associated with both fields and wave vectors perpendicular to the mean field. Cascades are more effective perpendicular to the mean field, since parallel fluctuations must bend the field lines. Thus, it is unlikely that space plasmas have isotropic fluctuations, and spectra may vary in different directions (recent studies suggest such anisotropy). The results of further study of symmetry will be important for understanding cosmic-ray modulation and solar energetic particle propagation.

Coherent cross-scale coupling occurs in a variety of regions within the heliosphere. Important questions concerning coherent coupling processes such as reconnection at Earth's magnetopause are posed in Chapter 2. Significant questions remain, too, on the origin, evolution, and role of turbulence in space plasmas. Often the most promising approach to the study of turbulence is simulation coupled with the observation of real plasmas. Various models of turbulence also show promise, although they must be continually checked against simulation and observation for success. Answering these questions will enhance our understanding of turbulence in general, thus increasing our ability to apply our ideas to astrophysical situations where direct measurements are not possible.

Outstanding Questions About Cross-Scale Coupling and Turbulence

- What is the detailed structure of heliospheric turbulence, and what are the consequences of this structure for the propagation of energetic particles?
- To what extent is observed turbulence actively cascading as opposed to being a "fossil" record of prior nonlinear processes?
- Does turbulent heating play a significant role in any of the main areas that have been suggested, namely the corona, the heliosphere, and Earth's magnetosphere?
- What is the mechanism for the dissipation of turbulence?

UNIVERSALITY OF STRUCTURES AND TRANSIENTS

Structures and transients are ubiquitous in the observable universe and play key roles in the redistribution of energy and momentum. The challenge is to determine the processes responsible for generating them. Further, it is important to understand how these processes are modified by vastly different scales and external boundary conditions in astrophysical settings. Astrophysical shocks are an example of this challenge. Earth's bow shock and a shock at a supernova remnant are distinctly different structures. However, understanding how interplanetary shocks differ from Earth's bow shock and how the heliospheric termination shock differs from interplanetary shocks leads to an understanding of how shock structure is modified when the pressure is dominated by a very energetic particle population. This understanding can then be extended to extreme cases such as shocks at supernova remnants.

There are other structures where analogies may reveal themselves through joint study. Do the fine tendrils observed within the Crab Nebula carry electric current, and are they analogous to the current filaments responsible for generating the aurora? Are the bipolar jets that are thought to help carry away angular momentum from collapsing protostars analogous to the auroral discharges at Jupiter, similarly responsible for shedding angular momentum (albeit small amounts) from that spinning body?

There are also important analogies between turbulence in the solar wind and in the diffuse interstellar medium. Resolving important questions concerning turbulence in the solar wind will provide important insight into the properties of the interstellar medium. Extending our understanding of turbulence in these regimes to other environments, such as star-forming regions in dense molecular clouds, is also important. In these regions, it appears that the turbulent energy decay rates are found to scale much the same way as in the heliosphere. This turbulence may play a role in stochastic acceleration of charged particles, becoming a possible mechanism of reacceleration of cosmic rays in the Galaxy.

The universal tendency of plasmas to couple across scales is exemplified by magnetic "islands" found in structures as different as astrophysical jets and the boundary layers of Earth's magnetosphere. Some aspects of such islands, as seen in the optical image of Quasar 3C 273 (Figure 3.9), have been attributed in similar jets to driven Kelvin-Helmholtz instabilities, while the x-ray emissions have been associated

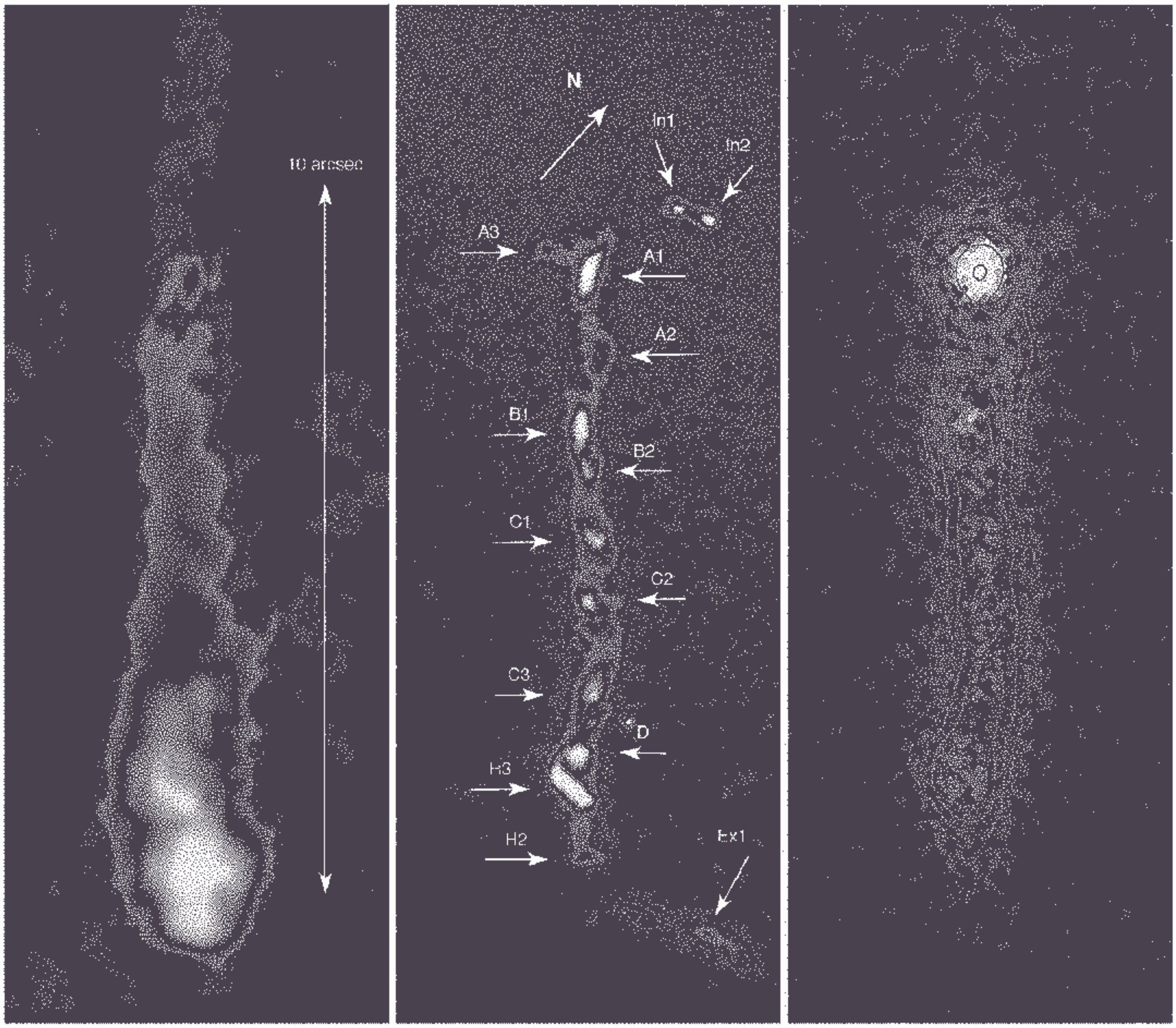


FIGURE 3.9 Images of an astrophysical jet from Quasar 3C 273 in ground-based radio (1.647 GHz), Hubble optical (617.0 nm), and Chandra x-ray (with optical overlay) bands. Features are labeled in the Hubble image as noted by J.N. Bahcall, S. Kirhakos, D.P. Schneider, R.J. Davis, T.W.B. Muxlow, S.T. Garrington, R.G. Conway, and S.C. Unwin, HST and MERLIN observations of the jet in 3C273, *Astrophysical Journal* 452, L91-L93, 1995. Courtesy of H. Marshall (Massachusetts Institute of Technology). Reprinted, with permission, from H.L. Marshall et al., Structure of the x-ray emission from the jet of 3C 273, *Astrophysical Journal* 549(2), L167-L171, 2001. Copyright 2001, American Astronomical Society.

with shocks that form as the supersonic jet is slowed by interactions with the ambient medium. The x-ray emissions from the jet islands could also be caused by particles accelerated in magnetic reconnection events. Similar island formation occurs in plasma boundary layers in the terrestrial magnetosphere (Figure 3.10). These islands can be caused by Kelvin-Helmholtz MHD instabilities or by ion-tearing instabili-

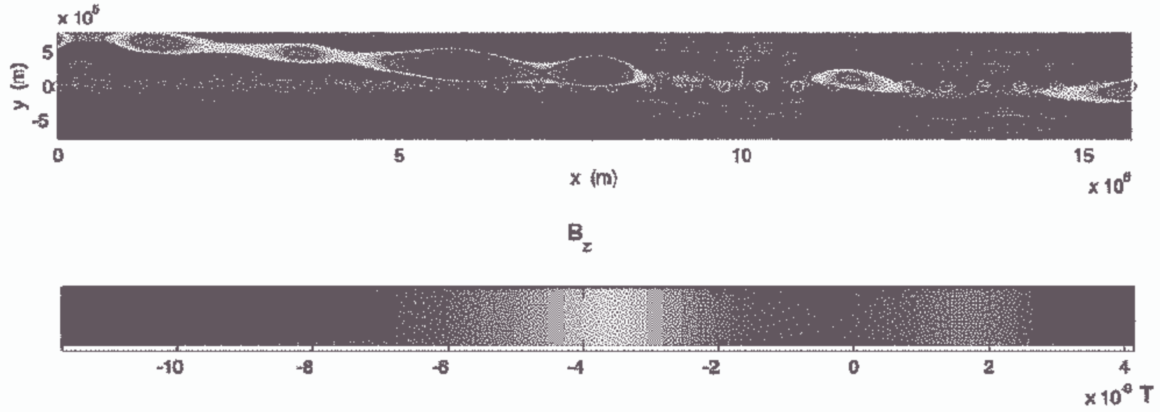


FIGURE 3.10 Magnetic islands along Earth's magnetopause deduced from in situ magnetic field data. These islands have been attributed to the ion-tearing mode. Reprinted, with permission, from L.-N. Hau and B.U.Ö. Sonnerup, Two-dimensional coherent structures in the magnetopause: Recovery of static equilibria from single-spacecraft data, *Journal of Geophysical Research* 104(A4), 6899-6918, 1999. Copyright 1999, American Geophysical Union.

ties and the associated magnetic reconnection. As these examples show, magnetic island structures resulting from high-speed differential plasma flows are present in both astrophysical and heliospheric environments, although at vastly different parameter scales. Even so, common phenomena such as MHD instabilities, shocks, and magnetic reconnection have been invoked to explain the structures in both environments.

Despite the vast differences in parameter regimes and boundary conditions that distinguish solar system and astrophysical plasma structures, the underlying plasma physical processes that give rise to and power these structures are the same. Ultimately, a fuller understanding of these processes in a general sense should be obtained with contributions from both space physics and plasma astrophysics.

NOTES

1. Tutorial articles on collisionless shocks and review articles on shock research in the first half of the 1980s can be found in R.G. Stone and B.T. Tsurutani, *Collisionless Shocks in the Heliosphere: A Tutorial Review* and *Collisionless Shocks in the Heliosphere: Reviews of Current Research*, Geophysical Monographs 34 and 35, American Geophysical Union, Washington, D.C., 1985. Developments in collisionless shock research since the mid-1980s are reviewed by T. Onsager and M.F. Thomsen, in *Reviews of Geophysics*, Supplement, 1991, and N. Omidi, *Reviews of Geophysics*, Supplement, 1995.
2. Cf. the discussion of shock acceleration in Chapter 6.
3. G.-C. Fälthammar, S.-I. Akasofu, and H. Alfvén, The significance of magnetospheric research for progress in astrophysics, *Nature* 275, 185-188, 1978.
4. Research on the heliosphere and its complex interface with the local interstellar medium is reviewed by T.E. Holzer, Interaction between the solar wind and the interstellar medium, *Annual Review of Astronomy and Astrophysics* 27, 199-234, 1989; and S.T. Suess, The heliopause, *Reviews of Geophysics* 28, 97-115, 1990.
5. B.F. Wood and J.L. Linsky, The local ISM and its interaction with the winds of nearby late-type stars, *Astrophysical Journal* 492, 788, 1998.
6. T. Gold and F. Hoyle, On the origin of solar flares, *Monthly Notices of the Royal Astronomical Society* 120, 89, 1960.
7. G.H. Fisher, The solar dynamo and emerging flux (invited review), *Solar Physics* 192, 119-139, 2000.

8. L. Burlaga, Global configuration of a magnetic cloud, pp. 373-377 in *Physics of Magnetic Flux Ropes*, Geophysical Monograph 58, C.T. Russell, E.R. Priest, and L.C. Lee, eds., American Geophysical Union, Washington, D.C., 1990.

9. C.T. Russell and R.C. Elphic, Observation of magnetic flux ropes in the Venus ionosphere, *Nature* 279, 616-618, 1979.

10. D.G. Sibeck et al., Magnetotail flux ropes, *Geophysical Research Letters* 11, 1090-1093, 1984.

11. See M. Hesse and M.G. Kivelson, The formation and structure of flux ropes in the magnetotail, in *New Perspectives on the Earth's Magnetotail*, Geophysical Monograph 105, A. Nishida, D.N. Baker, and S.W.H. Cowley, eds., American Geophysical Union, Washington, D.C., 1998.

12. M.L. Goldstein, D.A. Roberts, and W.H. Matthaeus, MHD turbulence in the solar wind, *Annual Review of Astronomy and Astrophysics* 33, 283-326, 1995.

13. A.N. Kolmogoroff, The local structure of turbulence in incompressible viscous fluids for very large Reynolds numbers. *C. R. Acad. Sci. URSS* 30, 376-387, 1941.

4

Plasma Interactions

Plasma populations throughout the universe interact with solid bodies, gases, magnetic fields, electromagnetic radiation, magnetohydrodynamic waves, shock waves, and other plasma populations. These interactions can occur locally as well as on very large scales between objects such as galaxies, stars, and planets. They can be loosely classified into electromagnetic interactions, flow-object interactions, plasma-neutral interactions, and radiation-plasma interactions.

Magnetic field lines connecting different plasma populations act as channels for the transport of plasmas, currents, electric fields, and waves between the two environments. In this way, the two plasmas become coupled electromagnetically to one another. Examples of *electromagnetic interactions* include the transfer of mass, momentum, and energy between Earth's magnetosphere and ionosphere; the outward transport of angular momentum in the jovian magnetosphere; and the production of accretion disks around protostars.

When a flowing magnetized plasma strikes a solid object, an atmosphere, or a magnetosphere, strong interactions of various types can occur. *Flow-object interactions* range from the simple sputtering of ions from solid surfaces (like the Moon) to the production of flux ropes around unmagnetized planets with atmospheres (like Venus), to magnetic reconnection and the resulting production of large-scale disturbances (like magnetic storms) at planets with magnetospheres.

Throughout the solar system and universe, plasmas are generally embedded in a background neutral gas with which they interact. *Plasma-neutral interactions* range from ion drag and "flywheel" effects in collision-dominated ionospheres; to charge-exchange reactions in the rarefied plasmas of magnetospheres and stellar winds; to dust-plasma interactions in cometary atmospheres, interstellar molecular clouds, protoplanetary disks, planetary rings, and stellar nebulas.

Radiation-plasma interactions are important in solar and stellar atmospheres, which respond to and mediate radiation in the form of magnetohydrodynamic waves and shocks emanating from the stellar surfaces and more energetic ultraviolet and x-ray photons propagating downward from the stellar coronas. These interactions will determine, for example, how ultraviolet emissions observed from stellar atmospheres are best interpreted in terms of their vertical structure.

Finally, the interactions described here take place over tremendous ranges of temporal and spatial scales. The spatial scales are often classified in terms of microscales (at which individual particle motions are important), mesoscales (which exhibit plasma fluid effects), and macroscales (comprising large structures such as coronal mass ejections and entire magnetospheres). Often the mesoscale and macroscale dynamics are produced by microscale phenomena (as magnetic reconnection leads to coronal mass ejections and magnetospheric substorms), while macroscale phenomena can drive dynamics at the smaller scales (as the Kelvin-Helmholtz instability is generated by large-scale flows of plasma along a boundary layer). As discussed in the preceding chapter, it is a fundamental property of space and astrophysical plasmas that efficient communication can occur across the various spatial scales.

In the following sections, the various plasma coupling phenomena are described briefly and some of their universal aspects are noted. Throughout there are close connections to material addressed, for example, in Chapters 2, 3, and 5.

ELECTROMAGNETIC INTERACTIONS

The coupling of different spatial domains along extended magnetic field lines can occur via field-aligned particle flows, electric fields, currents, and parallel propagating waves. At Earth, the most important manifestation of this process is the strong electromagnetic coupling that occurs between the magnetosphere and the ionosphere.

This coupling includes plasma circulation, plasma escape along field lines, field-aligned particle acceleration, and parallel (to **B**) currents. The current density along magnetic field lines is provided by electrons from the ionosphere (the downward currents) and the much more tenuous magnetosphere (the upward currents). Since the magnetospheric densities are low (a few per cubic centimeter at most at Earth), intense upward currents require field-aligned electric fields, which accelerate magnetospheric electrons down into the atmosphere to produce the required current and, in the process, create bright auroral forms. Figure 4.1 shows a view of Earth's northern aurora from a spacecraft high overhead. The aurora consists of two components: the diffuse aurora, which covers a broad latitude range and is relatively structureless, and the highly structured and dynamic discrete aurora, whose bright forms are easily seen from the ground. The diffuse aurora is produced by particles precipitated from the near-Earth plasma sheet by wave-particle interactions, while the discrete auroral emissions are excited by beams of energetic electrons from the outer magnetosphere that have been accelerated in field-aligned electric fields and, in particularly dynamic situations or regions, by high-powered Alfvén waves (cf. Chapter 6).

The processes that drive field-aligned currents into ionospheric plasmas also generate electric fields transverse to the magnetic field, the strength and location of which are strongly influenced by the properties of the ionospheric plasma. A two-way coupling between such regimes is set up in response to the driving field-aligned currents. These transverse electric fields drive ionospheric circulation and, through ion-neutral collisions, the motion of the neutral atmospheric gas. Similarly, the ionospheric feedback electric fields map upward along the magnetic fields, affecting processes in the overlying regions. In the terrestrial environs, the development of a disturbance ring current drives strong electric fields in regions of low ionospheric conductivity. These, in turn, affect both the planet's thermal plasma envelope and the further development of the energetic-plasma ring current.

The electromagnetic coupling processes in the magnetosphere-ionosphere systems of other planets are much less well understood than Earth's, but they offer an elegant array of plasma dynamical processes. Jupiter's magnetosphere is an especially rich environment for testing theories about electromagnetic coupling. In contrast to Earth's magnetosphere, where the dynamics are driven by energy extracted from the solar wind interaction, the jovian magnetosphere is powered by the planet's rotational energy that is transferred to the magnetosphere by field-aligned currents that couple the ionosphere with the magnetospheric plasma and set the plasma into corotational motion. Jupiter is of special astrophysical interest

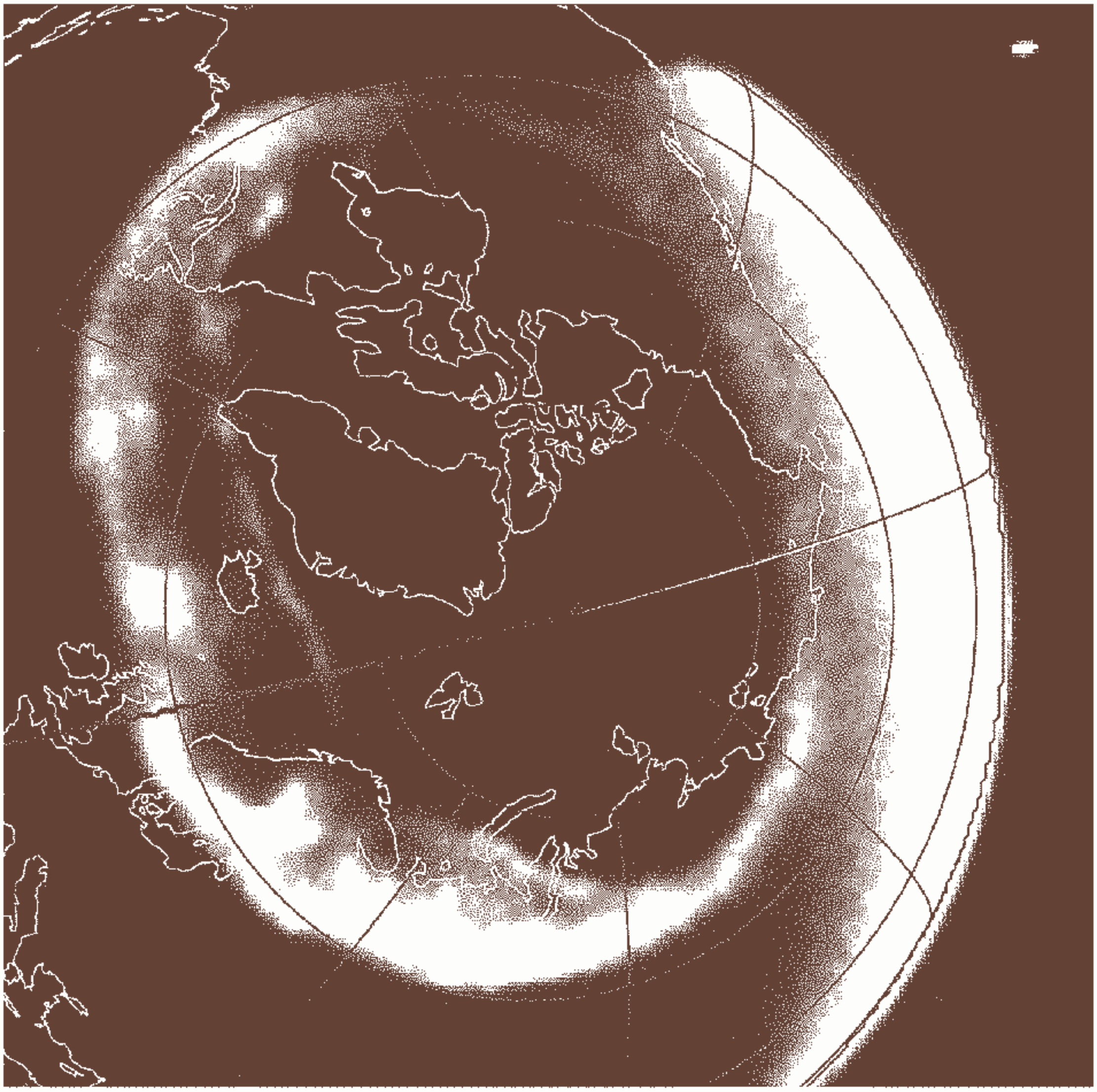


FIGURE 4.1 Earth's northern aurora, as viewed with the far-ultraviolet camera on the IMAGE spacecraft. Auroral emissions can be seen extending to high latitudes on the nightside. Also evident are the terminator (the boundary between the dayside and the nightside) and, at the right of the image, bright ultraviolet dayglow emissions from the sunlit hemisphere. Courtesy of NASA and the IMAGE Far-Ultraviolet Imaging Team.

---

because the transfer of torque through the electromagnetic coupling of the planet to the magnetosphere is a close analogue to the shedding of angular momentum from a central body to a surrounding nebula by means of magnetic fields and field-aligned currents thought to occur in other astrophysical environments (see sidebar, "The Formation of Stellar and Planetary Systems").

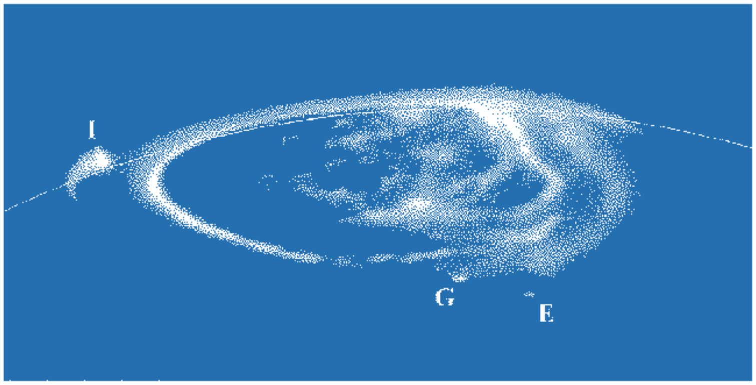


FIGURE 4.2 Jupiter's northern aurora, as viewed at ultraviolet wavelengths with the Hubble Space Telescope. The image shows the main auroral oval, diffuse polar cap emissions, and auroral emissions at the magnetic footprints of Io (I), Ganymede (G), and Europa (E). Image courtesy of J. Clarke (Boston University). Reprinted, with permission, from J.T. Clarke et al., Ultraviolet emissions from the magnetic footprints of Io, Ganymede and Europa on Jupiter, *Nature* 415 997-1000, 2002. Copyright 2002, Macmillan Publishers Ltd.

A distinctive feature of Jupiter's magnetosphere-ionosphere system is the fact that the interaction of the Galilean moons with the magnetospheric plasma generates electrical currents that couple the moons to Jupiter's ionosphere. The electrodynamical interaction between Jupiter and Io has been known for some time and is evidenced by radio emissions and auroral emissions at the foot of the flux tube linking the planet to the satellite. Recent Hubble Space Telescope observations of similar localized emissions at the magnetic footpoints of Ganymede and Europa are evidence that these moons, too, are electrically coupled by field-aligned currents with the jovian ionosphere (Figure 4.2).

FLOW-OBJECT INTERACTIONS

When a flowing magnetized plasma encounters an obstacle, such as another magnetized or a nonmagnetized plasma, relatively sharp boundaries tend to form that act to separate the plasmas. In the case of the supersonic solar wind encountering a planetary obstacle, the outermost boundary is a bow shock (discussed in Chapter 3), which heats and slows down the solar wind so that it can flow around the obstacle. If the planet is strongly magnetized, the solar wind is separated from the planetary plasma environment by a boundary known as a magnetopause. Magnetopauses exist at Earth, Jupiter, Saturn, Uranus, Neptune, and Mercury. The volume of space inside the magnetopause is dominated by the plasma and magnetic field associated with the planet, while outside the magnetopause the solar wind's plasma and magnetic field dominate. A magnetopause also exists at the jovian moon Ganymede, whose intrinsic magnetic field was discovered in 1996; in this case, however, the ambient plasma is that of the jovian magnetosphere rather than the solar wind.

The separation of solar wind and magnetospheric plasmas is not perfect, owing to a dissipation of the magnetopause current, allowing plasma and electric fields to penetrate the magnetopause. At Earth, this interaction produces a dynamic response that depends to a certain extent on the properties of the upper atmosphere, producing heat and auroral light emissions. Since Mercury, the only other terrestrial planet with a significant global magnetic field, has no atmosphere, but only a tenuous exosphere of sodium and

THE FORMATION OF STELLAR AND PLANETARY SYSTEMS

Angular momentum shedding from a centralized spinning region to a surrounding nebula is thought to happen, for example, in the early phases of stellar and planetary system formation. The process of stellar collapse can be summarized through the following paradigm:[1] A protostellar core collapses inside out, and the initial angular momentum of the system produces an accretion disk. This disk transfers mass onto the central protostar while angular momentum is transferred outward. In general, it appears that the formation of a jet combined with an accretion disk is a crucial element of angular momentum shedding. These processes (in particular for high-mass stars) are poorly understood, since an adequate description of viscosity in hydromagnetic disks is still lacking. Nevertheless, the presence of a disk, as well as jets, has been observed and provides a mechanism for the formation of planetary bodies.

Many such mysteries surround the early phases of the formation of stellar and planetary systems. Among these are the following (see figure, p. 51): What processes control the collapse of molecular cloud cores during the earliest phases of stellar system formation? Strong coupling between the neutral gas and dust components and the magnetized plasma components is thought to play a major role. The process of ambipolar diffusion that allows these components to separate is not well understood. That process has close analogues with the plasma-neutral interactions (see discussion in this chapter) occurring near the heliospheric boundaries, the plasma-neutral coupling occurring in planetary upper atmospheres, and the plasma-netural-dust interactions occurring in the neighborhood of comets. Solar system plasma physics has much to contribute to this topic. A related question is, What role does hydromagnetic turbulence play in the initial cloud-core collapse? Similar issues of turbulent transport processes surround the solar-system analogues to this problem already mentioned in Chapter 3. Other questions include: How do protostars shed > 98 percent of their angular momentum as they collapse into stellar and planetary systems? How are bipolar jets created and maintained?

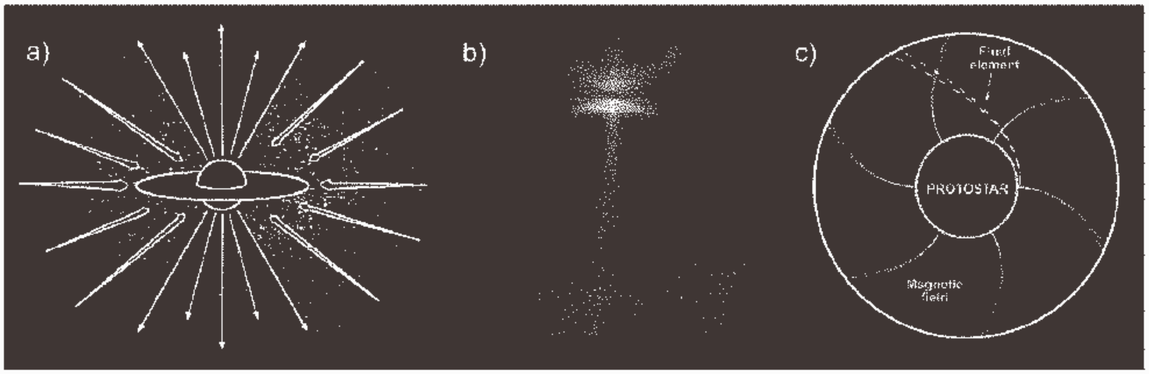
---

[1] Cf. F.J. Shu et al., The collapse of clouds and the formation and evolution of stars and disks, in *Protostars and Planets III*, University of Arizona Press, Tucson, 1993.

potassium, its magnetosphere probably responds in a fundamentally different way from Earth's magnetosphere to solar-wind variations. In fact, the role of the ionosphere in planetary magnetospheres can perhaps best be understood by exploring a magnetosphere, such as Mercury's, that has no ionosphere. The NASA Messenger mission, now under development, will take this important next step. The solar wind most certainly breaches, by magnetic reconnection, the magnetopauses of Jupiter and the other gas giants as well, but the extent of the contribution of the resulting energy transfer to magnetospheric dynamics at these planets is not known.[1]

In the case of nonmagnetized bodies, such as Venus, Mars, and comets, it is the planetary or cometary ionosphere, not a strong intrinsic magnetic field, that is the obstacle to the solar wind. The boundary that separates the solar wind plasma from the ionospheric plasma is called the ionopause (Figure 4.3). Unlike the ionospheric plasma, the body's neutral atmosphere is not confined by this boundary and extends beyond it into the solar wind-dominated region. Here, some of the neutral atoms or molecules are converted by photoionization, impact ionization, and charge exchange into ions, which are then picked up by the solar wind's motional electric field, mass loading the solar wind and slowing its flow. In addition to the thermal pressure of the ionosphere against the solar wind, the solar wind is also opposed by a magnetic barrier that

Some researchers have developed models of angular momentum shedding where the shedding occurs via magnetic field torquing. Magnetic field forcing also plays a central role in some theories of bipolar jet formation, which further aids the shedding of angular momentum. The physics involved in all of these applications is fundamentally similar to the physics in processes ongoing in solar system plasmas.



Bipolar outflows and the shedding of angular momentum during star formation. (a) Sketch illustrating the stage in stellar formation when matter from the molecular cloud core continues to accrete on the circumstellar disk while collimated jets have formed from both poles. (b) A Hubble Space Telescope image of the young stellar object HH 30 showing the bipolar outflows and the circumstellar dust disk. (c) A sketch of a protostar viewed in the equatorial plane illustrating the interaction between the protostellar magnetic field and the surrounding accretion disk by means of which angular momentum is transferred outward. Panel (a) is reprinted by permission from F.H. Shu et al., Star formation in molecular clouds: Observation and theory, *Annual Review of Astronomy and Astrophysics* 25, 23-81, 1987. Copyright 1987, Annual Reviews www.annualreviews.org. Panel (b) is courtesy of C. Burrows (Space Telescope Science Institute and the European Space Agency), the WFPC 2 Investigation Definition Team, and NASA. Panel (c) is reprinted, with permission, from J.R. Najita and F.H. Shu, Magnetocentrifugally driven flows from young stars and disks. III. Numerical solution of the sub-Alfvénic region, *Astrophysical Journal* 429, 808-825. Copyright by the American Astronomical Society.

forms because of the piling up of solar wind magnetic field lines at the ionopause as the solar wind plasma is slowed and compressed by the encounter with the ionospheric obstacle. As in the case of magnetized bodies, the separation of the solar wind plasma and the planetary or cometary plasma is not perfect, and at times of high solar wind dynamic pressure, the solar wind magnetic field may penetrate into the ionosphere.

Of the various heliospheric plasma interactions, the one between the solar wind and Earth's magnetospheric/ionospheric system has the most relevance for human activities and is by far the best studied. From a wide range of observations, it has been concluded that a small fraction of the solar wind mass, energy, and momentum incident upon Earth's magnetosphere is allowed to penetrate the magnetopause. Once inside the magnetosphere, this solar wind energy powers high-latitude ionospheric convection, generates field-aligned currents into and out of the ionosphere, initiates geomagnetic storms and substorms, produces the ring current, and drives auroral displays. All of these phenomena intensify during periods of southward interplanetary magnetic field (IMF) orientation.

Many mechanisms, including diffusion, impulsive penetration, and the Kelvin-Helmholtz instability, have been proposed to account for the interaction of the solar wind with the terrestrial magnetosphere. Only one, magnetic merging (or reconnection), predicts the observed relationships between the IMF

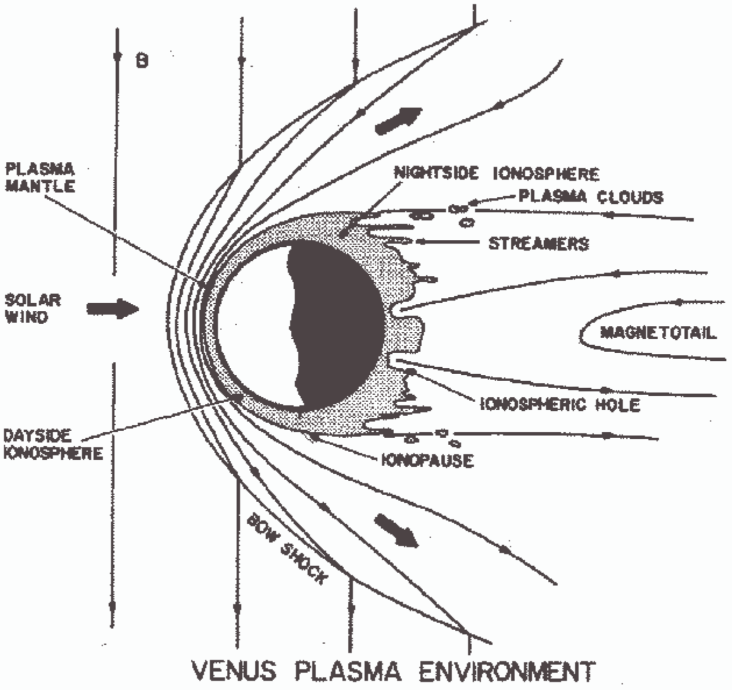


FIGURE 4.3 Schematic illustrating the interaction of the solar wind with Venus. Reprinted, with permission, from T.E. Cravens, The solar wind interaction with non-magnetic bodies and the role of small-scale structures, pp. 353-366 in *Solar System Plasma Physics*, Geophysical Monograph 54, J.H. Waite, Jr., J.L. Burch, and R.L. Moore, eds., American Geophysical Union, Washington, D.C., 1989. Copyright 1989, American Geophysical Union.

orientation and ionospheric and magnetospheric phenomena. However, there are many models for magnetic merging on Earth's magnetopause. Some models propose that merging always occurs in the vicinity of the subsolar point on the magnetopause, others that its location depends on the IMF orientation. Some models propose that it occurs steadily, others that it takes place in bursts. Some models suggest that bursty merging occurs in response to varying solar wind conditions, others that it occurs in response to intrinsic magnetopause instabilities. Some models require that it occurs along an extended line, others that it takes place in patches.[2]

In situ measurements over the past 20 years have revealed convincing evidence for merging at the magnetopause, namely, mixed magnetosheath and magnetospheric plasmas, accelerated plasma flows, magnetic field components normal to the nominal magnetopause, and streaming electron populations. Because almost all of these studies were based on single-point measurements during transient magnetopause crossings, they could not determine the extent of merging, its duration, whether it was more rapid in the subsolar region or elsewhere, or whether it was triggered by varying solar wind conditions. Recent imaging of the proton aurora by the NASA IMAGE satellite, which can identify protons accelerated by the reconnection electric field as they bombard the dayside upper atmosphere, has shown that magnetic reconnection occurs continuously at the magnetopause, changing location in response to variations in the direction of the solar-wind magnetic field. In situ measurements by the four-spacecraft Cluster II mission have confirmed that the reconnection regions connect to the proton aurora emission regions.[3]

PLASMA-NEUTRAL INTERACTIONS

Plasma-neutral interactions involve the transfer of charge, momentum, and energy in ion-neutral and electron-neutral collisions. Important examples are the resonant charge exchange interaction between an ion and its parent neutral ($H^+ + H \Leftrightarrow H + H^+$) and the accidentally resonant charge exchange reaction $O^+ + H \Leftrightarrow H^+ + O$. In planets with radiation belts, the trapped energetic H^+ and O^+ ions can charge exchange with the background hydrogen gas, and in this way an energetic trapped ion becomes an energetic escaping neutral atom (Figure 4.4). Such energetic neutral atoms, which retain the energy and velocity of the parent ions, can be detected remotely to produce global images of the magnetospheric ion populations. Near ionopauses, the relatively hot solar wind ions can exchange charge with planetary neutrals, thereby affecting the plasma populations there. Exchange of charge with solar wind protons is one of the mechanisms by which inflowing interstellar neutrals are converted to ions within the heliosphere. (The other mechanism is photoionization by solar ultraviolet radiation.) The ions newly created by charge exchange and photoionization are picked up by the solar wind and transported outward, toward the termination shock, where some of them are accelerated to extremely high energies. These return to the inner heliosphere as anomalous cosmic rays. Charge exchange also plays a major role in establishing the structure of, and

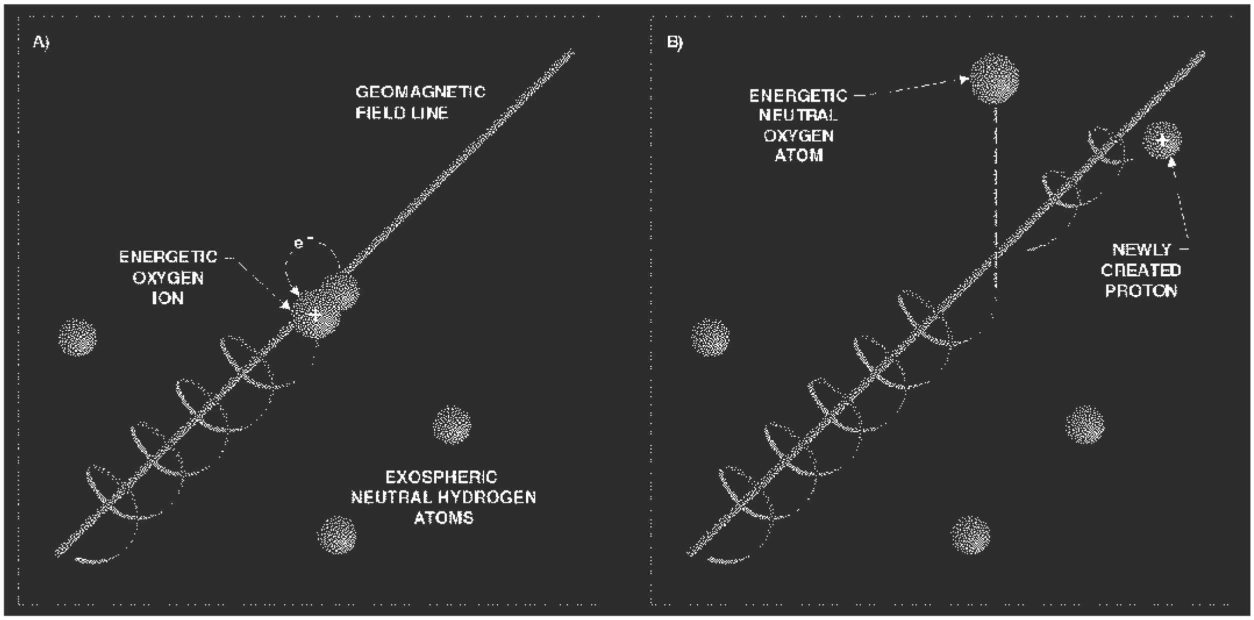


FIGURE 4.4 The creation of an energetic neutral atom through charge exchange. An oxygen ion trapped in the geomagnetic field captures an electron from a hydrogen atom in Earth's extended neutral atmosphere (exosphere) (A). The resulting energetic neutral oxygen atom is no longer trapped and can travel in a line-of-sight path away from the source population (B). The detection of such energetic neutral atoms by a remote imager allows global images to be made of magnetospheric plasmas, which are invisible to standard astronomical observing techniques.

populations within, the boundary regions of our heliosphere. It is, for example, responsible for the neutral hydrogen wall between the termination shock and the heliopause.

In addition to charge exchange, other ion-neutral collisional processes can affect the momentum and energy transfer between different spatial domains, such as between ionospheres and magnetospheres. In the upper atmosphere of planets or comets where significant neutrals exist, plasma dynamics both drives and responds to the neutral circulation, leading to coupling and feedback between these regions. In the case of Earth's ionosphere, ion convection driven by the interaction with the solar wind is generally faster than the motion of the ambient neutral gases. Neutrals are driven to move in the same direction as the ions due to ion drag or Ampere's force. When the reconnection rate at the dayside magnetopause is significantly reduced—for example, when the IMF suddenly turns from southward to northward—the magnetospheric driver of the ion motion is quickly reduced, whereas the neutrals tend to maintain their original inertial motion, forming a so-called fly-wheel effect. Under such circumstances, neutrals transfer energy and momentum to the ions, thus providing a mechanical and electromagnetic coupling from the thermosphere to the ionosphere and the magnetosphere.

RADIATION-PLASMA INTERACTIONS

Most of the information researchers have about astronomical objects comes from electromagnetic waves that are generated in and modified by the objects' dynamic gaseous envelopes or atmospheres. While radiative transfer in dynamic gaseous media is a well-developed discipline, the importance of the interaction between electromagnetic radiation and matter in the plasma state has only recently been recognized and analyzed. The Sun's chromosphere, photosphere, and corona represent a unique laboratory for studies of the production, transport, and absorption of electromagnetic radiation in a plasma. The results of such studies are relevant to our understanding of radiation-plasma coupling in other astrophysical systems and to the interpretation of electromagnetic emissions from remote astrophysical objects.

The atmospheric layers of the Sun are affected strongly by radiation-plasma coupling. The optical surface of the Sun, the photosphere, is an approximately 6000°C black body. Subsurface acoustic, gravity, and magnetohydrodynamic waves propagate upward from the photosphere and steepen into shocks as they rise through the overlying 20,000°C chromosphere (see Figure 4.5) because the gas and plasma densities decrease rapidly with altitude. The interaction of these waves and shocks with the chromospheric gas produces heat and ionization. At the same time, high-energy radiation in the form of ultraviolet and x-rays from the million-degree corona propagates downward through the transition region and into the chromosphere, producing additional ionization. These radiation-plasma interactions produce a temperature-altitude profile that is quite different from the profile predicted by the standard quasi-static models.[4]

The ionization profiles produced by the high-energy coronal radiation determine which particles are picked up most easily into the solar or stellar wind, and there is a well-known fractionation of solar and some stellar atmospheres in which elements with low first ionization potential (<10 eV) are enhanced over those with high first ionization potential. By the relative ion abundances in the solar wind it should be possible in principle to identify the source region of the wind, for example, in the chromosphere and to obtain information on heating mechanisms and magnetic topology in the source region.[5] This information will in turn provide a context for the interpretation of observations of other stellar coronas.

SUMMARY

Plasmas throughout the universe are strongly affected by the presence of magnetic fields and the currents that flow in response to any stresses placed on the magnetic field. Magnetized plasmas can

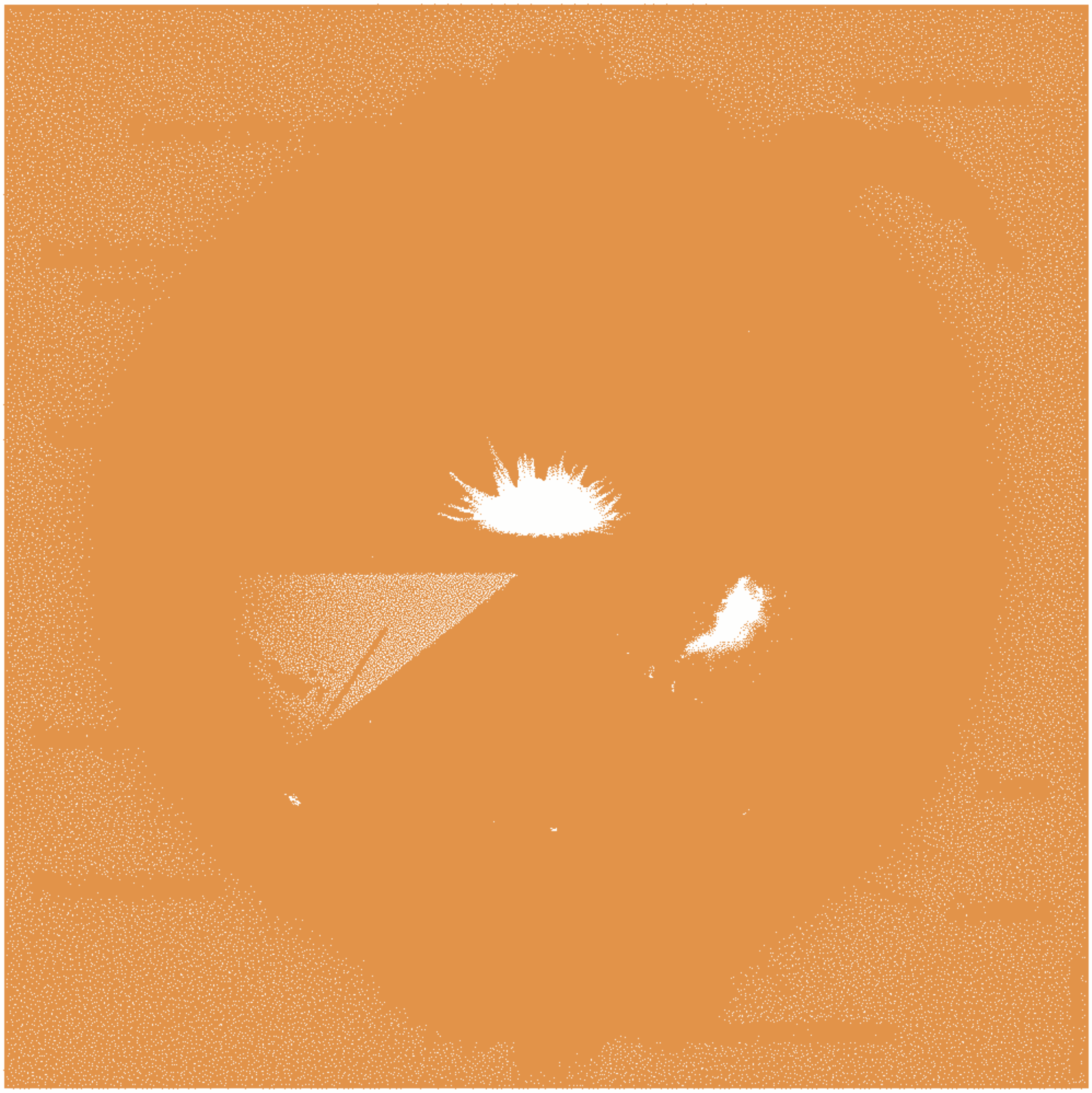


FIGURE 4.5 Cutaway drawing of the Sun showing some of its major regions and features. Courtesy of Steele Hill/NASA.

interact with their local environment, producing phenomena such as angular momentum shedding, which is considered to be an important mechanism for astrophysical processes such as protostar collapse, formation of accretion disks, and formation of planets. Plasma coupling can also occur over vast distances among completely different plasma environments. For example, the Sun is magnetically coupled to the planets and moons in the solar system; and some planets, such as Jupiter, are magnetically coupled to their moons. Magnetic fields provide the connection between different plasma environments that acts to introduce nonlocal characteristics into a local plasma environment. When the interacting plasmas do not represent hard obstacles to each other, the coupling along magnetic field lines is gradual and is characterized by large spatial scales. But if a flowing magnetized plasma encounters an obstacle, such as a planet with a strong intrinsic magnetic field, relatively sharp boundaries tend to form and in this case the magnetic coupling is characterized by small spatial scales. Such small-scale coupling occurs at bow shocks, magnetopauses, and ionopauses. At sharp boundaries, like these, a host of microphysical plasma processes can occur, including magnetic reconnection, particle acceleration, wave excitation, and the generation of parallel electric fields.

Phenomena associated with plasma-neutral interactions and radiation-plasma interactions mediate plasma coupling and can even represent controlling factors in the formation of plasma boundaries and the generation of stellar winds.

It is fair to say that the wide range of plasma interactions that occur in solar system and astrophysical settings is appreciated but not very well understood. Phenomena such as magnetic reconnection are known to be crucial, but the underlying mechanisms have not been experimentally verified. The list of outstanding questions concerning plasma interactions is quite long, but the prospects for their resolution have improved greatly because of the rapid development of numerical modeling techniques and the advances in the remote sensing and multipoint measurements of plasmas.

Outstanding Questions About Plasma Interactions

- How does the solar wind interact with a magnetized planet or moon that does not contain an ionosphere?
- Where, when, and how does magnetic reconnection occur at Earth's magnetopause?
- What is the source region of the solar wind in the chromosphere, and what are the source region heating mechanisms and magnetic topology?
- What is the ultimate cause of solar wind fractionation, and why are high first-ionization-potential ions more prevalent in the slow solar wind?
- What is the role of charge exchange in the coupling of the solar wind with magnetized and unmagnetized solar system bodies?
- What is the role of fluid turbulence in transporting mass and momentum across plasma boundary layers?
- How does micro-turbulence couple into mesoscale plasma dynamics?

NOTES

1. The most thoroughly studied of the outer-planet magnetospheres, Jupiter's, is powered primarily by planetary rotation. However, changes in solar wind ram pressure have been shown to modulate magnetospheric activity and magnetosphere-ionosphere coupling, and theoretical arguments have been put forward to explain certain auroral features as the ionospheric signatures of reconnection at Jupiter's dayside magnetopause (see S.W.H. Cowley, E.J. Bunce, and J.D. Nichols, Origins of Jupiter's main oval auroral emissions, *Journal of Geophysical Research* 108(A4), 8002, 2003; and S.W.H. Cowley, E.J. Bunce, T.S. Stallard, and S. Miller, Jupiter's polar ionospheric flows: Theoretical interpretation, *Geophysical Research Letters* 30(5), 1220, 2003).

2. Cf. the more detailed discussion in Chapter 2 of reconnection in Earth's magnetosphere.

3. T. Phan, H.U. Frey, S. Frey, L. Peticolas, S. Fuselier, C. Carlson, H. Rème, J.-M. Bosqued, A. Balogh, M. Dunlop, L. Kistler, C. Mouikis, I. Dandouras, J.-A. Sauvaud, S. Mende, J. McFadden, G. Parks, E. Moebius, B. Klecker, G. Paschmann, M. Fujimoto, S. Petrinec, M.F. Marcucci, A. Korth, and R. Lundin, Simultaneous Cluster and IMAGE observations of cusp reconnection and auroral proton spot for northward IMF, *Geophysical Research Letters* 30(10), 1509, 2003.

4. See, for example, M. Carlsson and R.F. Stein, Does a non-magnetic solar chromosphere exist?, *Astrophysical Journal Letters* 440, L29, 1995.

5. H. Peter, Element fractionation in the solar chromosphere driven by ionization-diffusion processes, *Astronomy and Astrophysics* 335, 691-702, 1998.

5

Explosive Energy Conversion

We owe our earliest awareness of magnetized cosmic plasmas to their propensity to explode. Throughout prehistory, magnetically mediated solar explosions (today called coronal mass ejections) occasionally lit the skies at night with auroras over the caves of our ancestors. But it was not until relatively recently, in the mid-19th century, that Richard Carrington, a solar astronomer, first witnessed a solar precursor—a solar flare—to an auroral night. Even earlier in the 19th century, Alexander von Humboldt first identified a terrestrial type of magnetically mediated explosion, which, to emphasize its explosive nature, he named a magnetic *storm.* (Humboldt recognized the episodic and, so, storm-like character of localized terrestrial magnetic disturbances that occurred suddenly around midnight. Researchers now refer to this type of localized magnetic disturbance as a "substorm" and use the term "magnetic storm" to refer to a global disturbance.) Also in the early 1800s, explosive auroral storms, now known to be auroral counterparts of Humboldt's localized magnetic storms, were recognized and described as "recurring *fits*." During the century following Carrington, solar flares and magnetic storms—both manifestations of magnetically mediated explosions, as mentioned—were topics at the center of interest of a new discipline that became known as solar-terrestrial relations.

After Sputnik, "solar-terrestrial relations" became "space physics," and space physicists, using data from spacecraft, began to expand their awareness of the explosive nature of magnetized cosmic plasmas. Terrestrial substorms, they found, are one of a hierarchy of explosive magnetospheric phenomena that begins with unnamed turbulence in the magnetotail, progresses through "bursty bulk flows" and "pseudo-breakups" to substorms and then to magnetic storms. It seems certain that the members of this hierarchy are related, but the relationships are unclear or controversial. Missions to planets other than Earth have added to the inventory of such phenomena. Mariner 10 recorded substorm-like events at Mercury, and Galileo in Jupiter's magnetosphere has observed dynamical events that appear to be related to substorms. Data from spacecraft have also expanded the known types of explosive solar phenomena. Beyond optical solar flares, newly identified types include coronal bright spots and x-ray flares. But the solar eruptive phenomenon most directly related to magnetic storms is the coronal mass ejection (CME), discovered in the mid-1970s in Skylab measurements (Figure 5.1).

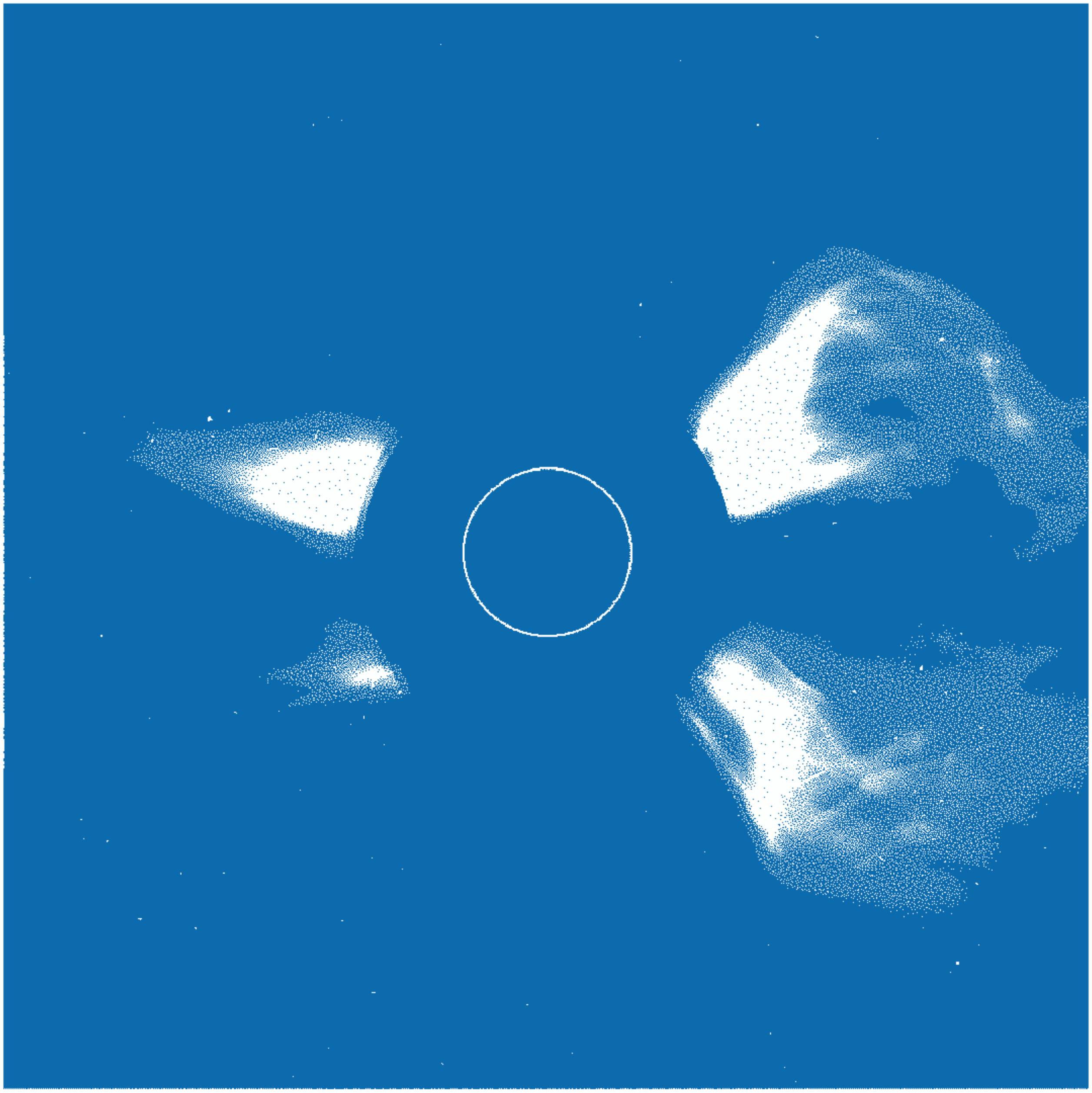


FIGURE 5.1 Large coronal mass ejection from November 6, 1997, as recorded by the LASCO C2 coronagraph at 12:36 universal time. The white inner circle represents the solar disk, which is hidden by the coronagraph's occulting disk (dark circle). Courtesy of SOHO (ESA and NASA).

These eruptive phenomena cover 13 orders of magnitude in energy and 5 orders of magnitude in time. Yet all are instances in which flow energy first converts gradually to magnetic energy and then explosively dissipates into kinetic energy and—in the case of flares—into electromagnetic radiation. A possible unifying concept that recurs in the following discussion is *storage-release*. That is, something stops magnetic energy from dissipating as fast as the flow generates it. So magnetic energy builds up until something causes the rapid dissipation of the stored energy to ensue. Most solar and magnetospheric theoretical research in this area concerns developing storing-and-releasing scenarios. The storage-release concept is deeply engrained in current thinking on explosive energy conversion, and much of the following discussion reflects its hegemony.

STORAGE-RELEASE IN THE SUN'S CORONA

A basic feature of the evolution leading to a solar explosive event such as a flare or coronal mass ejection is that magnetic energy is stored. The characteristic time scale for magnetic energy transfer through the solar surface, the photosphere, is much longer than the time scale for transfer through the corona. If the energy could be gradually released as it was introduced into the system, there would be no explosive energy release in the corona.

The underlying source of energy for all coronal activity is the mass motion in the subphotospheric convection region. Both the plasma beta and the magnetic Reynolds number[1] are much greater than unity in the photosphere and below and, consequently, the turbulent motions there tangle and stress magnetic field lines before and after the field emerges into the corona, so that the coronal field contains a large amount of free energy that can be released through magnetic reconnection processes that alter its topology. It is the free energy in these stressed coronal magnetic fields that powers (perhaps with intermediate steps) solar nonthermal emissions such as ultraviolet and x-ray radiation, solar wind outflows, and high-energy particles. Solar activity can be understood, therefore, as simply the transformation of the energy in mass motions in the Sun's convection zone into the energy in nonthermal emissions from the Sun's atmosphere, with the magnetic field acting as an intermediary.

It is not immediately obvious, however, why this energy transfer can lead to explosive phenomena such as CMEs or flares. The problem is that the Alfvén speed in the photosphere (<1 km/s) is at least three orders of magnitude smaller than the speed in the corona (>1000 km/s), which implies that the corona can easily adjust to changes in photospheric driving via a quasi-steady evolution. Indeed, this is usually the case. The time scales (of the order of days) for slow variations in the emissions from active regions, quiet regions, and coronal holes are commensurate with the slow evolution of the underlying photospheric magnetic field. As discussed in Chapter 2, however, reconnection based on classical resistivity is not fast enough to explain even these slow variations, much less explosive solar flares. The magnetic gradients must become steep enough, and the current sheets must become sufficiently intense, before fast reconnection can be triggered. Thus because some threshold must be reached before fast reconnection can occur, magnetic free energy can build to substantial levels before release. Solar flares, and indeed perhaps coronal heating, are then naturally storage-release mechanisms. It is less clear what the storage mechanism for CMEs is.

CME models can typically be sorted into three basic classes, using analogies to the dynamics of a spring (Figure 5.2). In the mass-loading model, chromospheric or coronal mass—for example, a prominence—weighs down a magnetic arcade or magnetic flux rope, stressing the magnetic spring. A CME occurs when the mass slips off, releasing the spring. Mass-loading models are concerned, in part, with specifying the slipping-off process. In the tether-release model a magnetic arcade, whose fields act like a set of tethers to constrain an underlying high-pressure magnetic configuration such as a flux rope, slowly weakens through magnetic reconnection. The final stage, before all magnetic tethers break, can occur

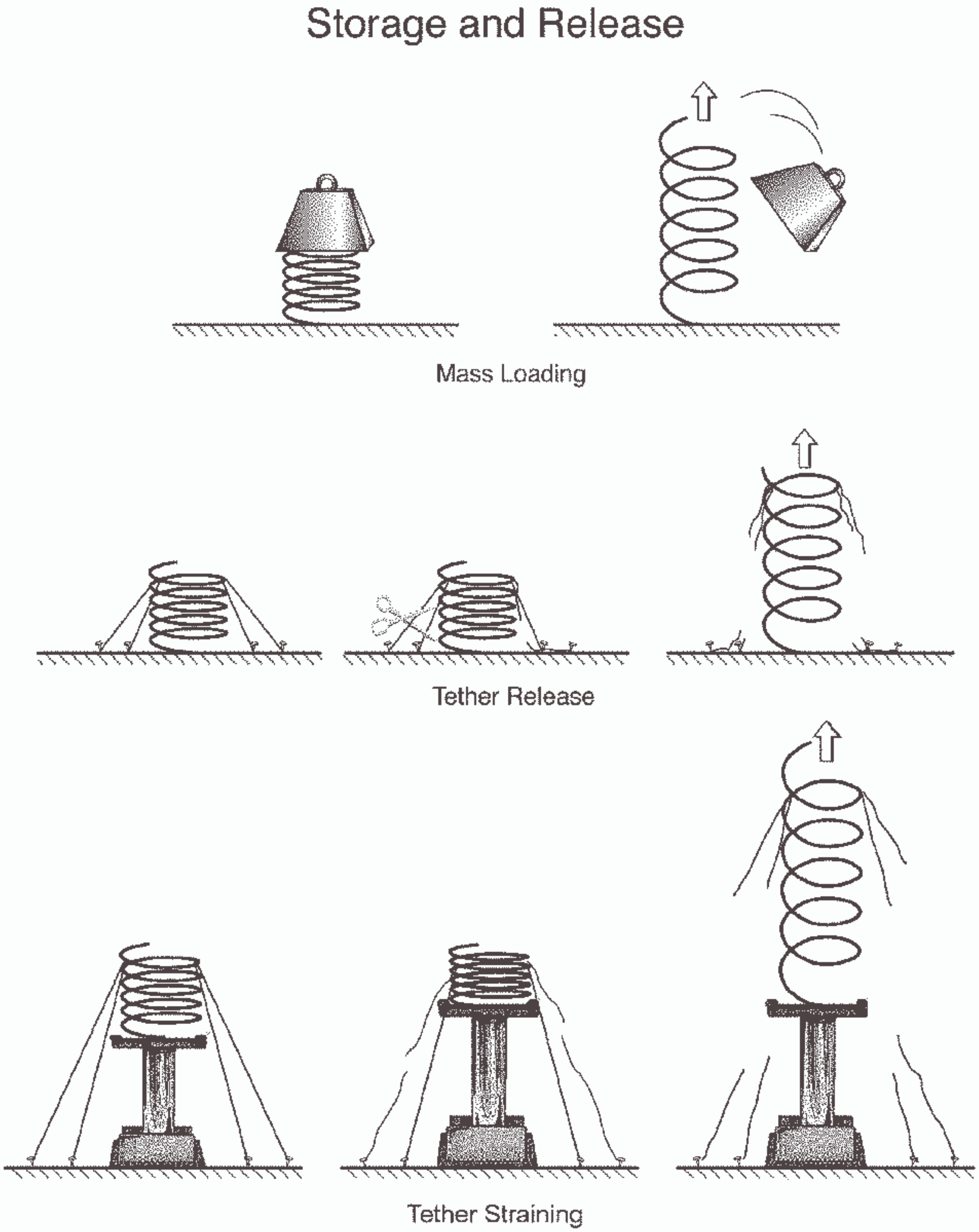


FIGURE 5.2 Three classes of coronal mass ejection models. Image courtesy of J. Klimchuk (Naval Research Laboratory). Reprinted, with permission, from J.A. Klimchuk, Theory of coronal mass ejections, pp. 143-157 in *Space Weather*, Geophysical Monograph 125, P. Song, H.J. Singer, and G.L. Siscoe, eds., American Geophysical Union, Washington, D.C., 2000. Copyright 2000, American Geophysical Union.

explosively, launching a CME. The third type of model, the tether-straining model, is a variation of the tether-release model. Here, the magnetic fields of an arcade are stressed through the growth of the underlying pressure. As in the case of tether release, the breaking of overlying magnetic fields through reconnection can occur explosively, releasing a CME. The mechanism for reconnection in the final phase, which gives the phenomenon its explosive character, is still a matter of conjecture. In every scenario that invokes some form of "tether cutting," the fast phase occurs either because some threshold is passed where reconnection switches from dormant to active or because a current sheet where reconnection can occur is suddenly created.

Since magnetic free energy powers all of these mechanisms, twisted or sheared magnetic field topologies, which can potentially change topology to release large amounts of energy, are required. The storage and release associated with CMEs can in fact be seen as a natural result of the conservation of magnetic helicity. Magnetic helicity is a property of the field related to its twist and linkage, and under coronal conditions the global helicity of a magnetic field is approximately conserved during reconnection. Thus reconnection by itself cannot release all of the free energy of a system, but only as much as can be released without altering the magnetic helicity of the system. In this manner, energy is built up in a twisted or sheared magnetic structure, until a mechanism such as those shown in Figure 5.2 leads to its eruption in a CME. The helicity of the structure is then bodily removed in the CME, and with its loss from the coronal system the rest of the stored energy can, in principle, be released.[2]

Outstanding Questions About Storage-Release in the Sun's Corona

- How is magnetic free energy built up and stored in the corona? For example, does it arise primarily from photospheric motions, or from the emergence of an already-twisted magnetic field from the solar interior?
- How is this magnetic free energy then converted into heating in solar flares and/or kinetic energy in coronal mass ejections?
- How significant is mass to the CME system?
- What is the magnetic topology of the solar corona before, during, and after a CME?

STORAGE-RELEASE IN EARTH'S MAGNETOTAIL

As discussed in Chapter 2, the magnetospheric substorm is the primary mode of magnetic energy conversion in the nightside magnetosphere. Observations clearly indicate that, at substorm onset, the magnetosphere suddenly and radically changes its structure within localized regions, its convection state, and its dissipation rate. Why does the magnetosphere fail to change smoothly from a state of slow convection and low dissipation to a state of fast convection and high dissipation?

A substorm has three main phases: a growth phase, during which magnetic energy is stored in the tail; an expansion phase, during which the stored energy is released; and a recovery phase, during which the magnetosphere returns to its pre-disturbance configuration (Figure 5.3). The growth phase begins when the interplanetary magnetic field (IMF) suddenly swings around to a southward orientation and reconnects abruptly with the northward geomagnetic field at the sunward magnetopause. As the open field lines are swept back into the magnetotail, magnetic flux is eroded from the dayside magnetopause and builds up rapidly in the northern and southern lobes of the magnetotail. The field intensity in the tail increases, and the plasma sheet (the reservoir of plasma between the tail lobes) thins. The growth phase continues until something triggers the explosive release of the accumulated magnetic energy in the expansion phase. In over half the substorms the expansion phase is triggered when the IMF turns northward again and dayside reconnection ceases.

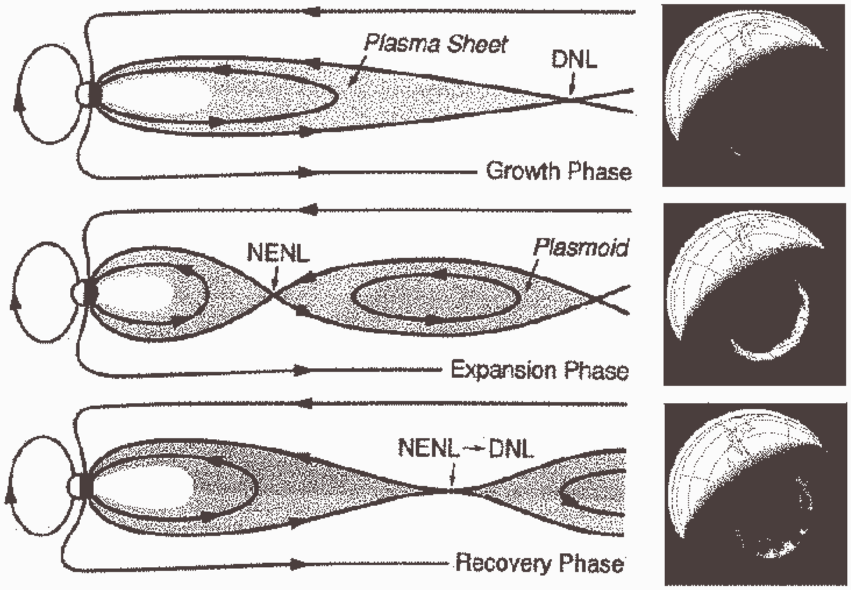


FIGURE 5.3 Schematic illustrating the three phases of a magnetospheric substorm, together with auroral images showing the auroral activity associated with each phase. The schematic is from W. Baumjohann and R.A. Treumann, *Basic Space Plasma Physics*, Imperial College Press, London, United Kingdom, 1996, and is reprinted by permission of the Imperial College Press. The auroral images were obtained with the IMAGE far-ultraviolet camera; IMAGE was located over the North Pole at an altitude of approximately 43,000 km.

During the expansion phase, the stretched lobe field lines reconnect earthward of the distant neutral line, which is formed by the merging of open field lines during quiet conditions, and form a new neutral line in the mid-tail, between 20 and 35 Earth radii. The stored magnetic energy is thereby converted into heat and plasma kinetic energy in the form of enhanced plasma flows and energetic particle injections. The field lines earthward of the newly formed neutral line assume a dipolar configuration and flow around toward the dayside, replenishing the magnetic flux that had been stripped away by magnetopause reconnection. The reconnected field lines tailward of the new neutral line form a closed magnetic structure known as a plasmoid, which is ejected down the tail at a speed of several hundred kilometers per second. As the substorm enters the recovery phase, the neutral line retreats from its location in the mid-tail and propagates down the tail, eventually becoming a new distant neutral line.

The basic substorm is quite a complicated transient affair, with specific auroral effects and magnetic fluctuations at the surface of Earth. In his pioneering paper on the auroral substorm,[3] Syun-Ichi Akasofu described the global auroral signature of the magnetospheric substorm and identified the brightening of an auroral arc in the midnight sector of the auroral oval with the onset of the substorm expansion phase. Following onset, the aurora intensifies and emissions move poleward of the oval, sometimes filling half the area of the polar cap. During the recovery phase, auroral activity decreases and auroral forms characteristic of the recovery phase, such as the double oval and eastward-drifting omega bands, are observed.

The details regarding how the stored magnetic energy in the tail lobes is transferred via tail reconnection to the plasma sheet and ultimately dissipated remain subjects for debate, and a number of different substorm models have been proposed. The two leading models are the near-Earth neutral-line (NENL) model, which is the one implicit in the substorm description above, and the current disruption model. In the NENL model, as the magnetic pressure and the currents intensify in the plasma sheet in the magnetotail, a threshold for the onset of rapid reconnection is exceeded, leading to explosive release of magnetic energy stored in the tail lobes and dipolarization of the tail magnetic configuration. In the current-disruption model, substorm breakup begins nearer Earth, at distances between 7 and 12 Earth radii, with the disruption of a very thin current sheet formed within the much thicker plasma sheet, while fast reconnection in the mid-tail develops later in the expansion phase. Observations have yielded conflicting evidence on whether the physical cause of breakup is in the near-Earth or mid-tail plasma sheet.

Outstanding Questions About Explosive Energy Release in Earth's Magnetosphere

- Why are substorms often triggered by a northward shift of the IMF after periods of southward IMF shift? What is the role of solar wind pressure changes?
- What is the nature of the threshold for the onset of rapid reconnection?
- What role does the ionosphere play in substorms?
- What are the nature and the importance of instabilities in the intense current layers in the near-Earth tail?

UNIVERSALITY OF STORAGE-RELEASE MECHANISMS

Solar active regions and planetary magnetospheres are typically driven by externally imposed forces that act over times that are long compared with the propagation time scale for magnetohydrodynamic waves. These systems then slowly evolve through a series of quasi-equilibrium states during which their magnetic field and plasma configurations are gradually driven far from possible minimum energy ground states. During this evolution, stress is accumulated as the system is prevented from returning to the ground state. Finally, these stressed configurations suddenly break, and the stored free energy is rapidly converted or dissipated into a variety of channels.

Plasma systems typically possess a wide variety of spatial and temporal scales. Sudden energy transfer events in these systems may have their ultimate origin in this multiscale property. Short space-time scale dynamics can break the local connection to the global constraining and driving forces and thereby open new channels of energy rearrangement and dissipation. The resulting changes take the system far from its initial stressed state and allow it to evolve to a lower-energy configuration. Conceptually, the change might be bimodal, with microscale changes affecting the macroscale, or may involve an even more complex turbulent multiscale system.

Thus the concept of storage and sudden release of energy is likely to be a universal one, naturally occurring in astrophysical systems possessing driving forces and multiple scales. Considered here in particular is how storage-release arises in magnetized plasmas, specifically in the context of the active Sun and the active magnetosphere of Earth. These solar system processes illustrate basic dynamical magnetic effects that occur throughout stellar systems, galaxies, and clusters of galaxies. We should be grateful for the simplicity of our local solar system laboratory, which has shown us so many effects but under conditions where their "simple" nature can be discerned.

The Sun, for all its magnetic complexity, is, after all, a relatively pedestrian star. One can only guess at the magnetic complexity of a Wolf-Rayet star. Multiple star systems, each star with its time-dependent magnetic fields and stellar wind, suggest a whole new level of magnetic complications. We would also expect more activity from a young star—if one thinks back to the young Sun when it had a rotation period of 2 days, the interplanetary magnetic field must have been very tightly wound, which in combination with the massive wind of its youth would no doubt lead to extremes of violent interplanetary dynamics. These statements are not mere speculations: evidence is beginning to accumulate from stellar observations. For example, a huge stellar eruption has been observed in the very young binary system of XZ-Tauri AB.[4] Eruptions have also been observed in a classic T-Tauri star and its stellar accretion disk that have been speculated to be akin to coronal mass ejections.[5] Moreover, the M-dwarfs, with their monstrous flares (which can be hundreds of times brighter than the brightest solar flares) and equally monstrous starspots (covering more than half the stellar diameter) also represent extreme applications of the basic principles discussed in this section. Whatever goes on in distant stellar and galactic systems, it involves the violent interaction of fields and plasmas, within and around stars, galactic nuclei, and their surrounding spaces.

NOTES

1. The plasma beta (β) is the ratio of the plasma pressure to the magnetic pressure. As discussed in Chapter 2, the magnetic Reynolds number (R_m) is the ratio of the magnetic diffusion time to the plasma flow time. When the magnetic Reynolds number is large, the magnetic field is "frozen" in the flow and moves with it.

2. B.C. Low, Solar activity and the corona, *Solar Physics* 167, 217, 1996. See also E.G. Blackman and A. Brandenburg, Doubly helical coronal ejections from dynamos and their role in sustaining the solar cycle, *Astrophysical Journal* 584, L99-L102, 2003.

3. S.-I. Akasofu, The development of the auroral substorm, *Planetary Space Science* 12, 273-282, 1964.

4. J.E. Krist, Hubble Space Telescope WFPC2 imaging of XZ Tauri: Time evolution of a Herbig-Haro bow shock, *Astrophysical Journal Letters* 515, L35-L38, 1999.

5. J.M. Oliveira, B.H. Foing, J.Th. van Loon, and Y.C. Unruh, Magnetospheric accretion and winds on the T Tauri star SU Aurigae: Multi-spectral line variability and cross-correlation analysis, *Astronomy and Astrophysics* 362, 615-627, 2000.

6

Energetic Particle Acceleration

As a consequence of the release of energy in a cosmic plasma, some portion of the background charged particle population is accelerated to very high—in some cases, relativistic—energies. Particle acceleration occurs throughout the universe, and the heliosphere provides the quintessential laboratory within which to investigate in situ the detailed character of different acceleration processes. Lessons learned here can often be translated to more exotic locations. Detailed models for particle acceleration have been developed by astrophysicists and space physicists for environments ranging from supernova remnant shock waves to flares on the Sun and stars to the magnetospheres of planets and pulsars. The development of these models provides one of the best examples of the cross-fertilization that can occur between space physics and astrophysics. For example, the theory of shock acceleration was originally developed in an astrophysical context. Its most refined development, however, has taken place in space physics owing to the availability of data from in situ observations of shocks in various solar system settings.

Cosmic acceleration processes can be grouped into three broad classes: (1) *shock acceleration*, (2) *coherent electric field acceleration*, and (3) *stochastic acceleration*. Both coherent and stochastic electric field acceleration can also occur as a part of the shock acceleration process. This chapter outlines the basic physical principles underlying each class of acceleration process and describes certain solar system plasma phenomena illustrative of the various processes.

SHOCK ACCELERATION

Shock Acceleration Mechanisms

One basic particle acceleration mechanism operating at shocks is known as diffusive shock acceleration or Fermi acceleration.[1] The workings of this mechanism are illustrated by the example of an elastic ball bouncing between two walls that are moving toward each other. In each collision with a wall, the ball not only changes direction but also increases its speed by a small increment proportional to the speed of the wall that it hits. No matter which wall the ball hits, its speed increases each time. This process will continue as long as the walls are moving together. Within a collisionless plasma, the reflecting "walls" are waves upstream and downstream from a shock wave. The waves are typically generated by the accelerated particles themselves. Particles may scatter off the upstream and downstream waves and bounce back and forth across the shock. If, as is usually the case, there is compression of the wave velocities at the shock, then the particles traversing the shock are accelerated, like the elastic ball between the approaching walls.

Diffusive shock acceleration occurs at both quasi-parallel and quasi-perpendicular shocks. ("Quasi-parallel" and "quasi-perpendicular" refer to the angle between the magnetic field and the shock normal. See the discussion of quasi-parallel and quasi-perpendicular shocks in Chapter 3.) At quasi-perpendicular shocks, in the absence of particle collisions with turbulence or waves, the compressed magnetic field downstream of a shock causes particles to drift along the shock face and to be accelerated in the upstream motional electric field. This coherent mechanism is referred to as shock drift acceleration.[2]

Heliospheric Shock Acceleration Sites

The solar wind flow speed is highly supersonic, and therefore shock waves will form ahead of any obstacle to the flow, or regions where high-speed plasma collides with low-speed plasma. The primary obstacles within the heliosphere are magnetic structures within the solar wind flow itself. The size of these obstacles, and the length of time during which particles can interact with them, determine the overall effectiveness of particle energization at these shocks. For example, at Earth's quasi-parallel bow shock, about 1 percent of the solar wind is accelerated from an initial energy of about 1 keV/e to energies of tens of keV/e. Shock acceleration at Jupiter's bow shock is proportionally larger because the interaction region is larger. The extreme example of this scaling (at least in the solar system) is at the heliospheric termination shock (the interface between the solar wind and the interplanetary medium where the solar wind is slowed to subsonic speeds). At this enormous shock, it is now thought that particles acquire energies of up to several hundred MeV/nucleon and become the anomalous cosmic-ray population. An additional consideration is the original seed population of ions that is accelerated, since this population is of varying character at different heliospheric sites. Below are enumerated the primary shock types in the heliosphere, each of which has an associated energetic particle population. The review starts at the Sun and moves outward to the heliosphere's termination shock.

Coronal Mass Ejections As the name implies, coronal mass ejections are events on the Sun wherein material from the corona (ranging from 10^{14} to a few times 10^{16} g) is ejected within a magnetic structure that moves out from the Sun at speeds of 10 to 2500 km/s. At solar maximum they occur as frequently as four to five times per day, and the fastest ones are associated with energetic particles. At solar minimum they rarely occur. At a typical speed of 600 km/s, a CME will reach the orbit of Earth about 3 days after it is launched from the Sun. As it moves out, it grows in size so that by the time it reaches Earth's orbit, it may be close to 1 AU across.[3]

CME-driven shocks can accelerate particles over basically the entire region from the Sun to Earth orbit and beyond. The shocks are shaped like quasi-spherical shells moving radially outward. They are broad and comparatively uniform so that the acceleration process, acting on particles moving along the shock front, can work maximally to completion before the shock finally fades away. Scientists know that significant acceleration occurs at the Sun itself because energetic particles propagate promptly to Earth well in advance of the arrival of the CME shock. Processes such as solar flares and coronal shocks often take place in association with CME eruptions. As the CME itself passes Earth, a further increase in energetic particles is often observed, and in this case is clearly associated with interplanetary acceleration near the CME. The seed population of particles available for acceleration by these CME shocks includes solar wind ions and other ions that may be present in the inner heliosphere; it may even include energized particles from earlier events. The range of speeds is larger for CME shocks than for any of the other heliospheric shocks, leading to a wide range of possible energies for the accelerated particles.

Corotating Interaction Regions The solar wind has a low-speed (~400 km/s) component, which originates from regions above closed magnetic loops, and a high-speed (~750 km/s) component, which originates from regions essentially free from overlying magnetic fields (coronal holes). Owing to the Sun's rotation, the solar wind streams move out into the interplanetary medium in a manner similar to water escaping from a rotating lawn sprinkler. The high-speed streams eventually overtake the low-speed streams, forming a compression region that is bounded by shock waves in the region beyond Earth's orbit. Because the solar wind streams corotate with the Sun, the regions of this fast/slow stream interaction are called corotating interaction regions (CIRs). Particles can interact with the CIR shocks for an extended period of time since the structure can be long-lived; in addition, the shock size is large, of the order of several astronomical units. Recent observational and theoretical studies have shown that, in addition to acceleration at the shocks bounding the compression region, particles are also accelerated by a Fermi-type process within the compression region itself.[4] CIR particles can achieve energies up to 10 to 20 MeV.

Planetary and Cometary Bow Shocks Planets with magnetic fields, such as Earth, Jupiter, Saturn, Uranus, and Neptune, have large standing shocks on the sunward side where the solar wind impacts the planet's magnetic field. Upstream of these shocks, energetic particle bursts are routinely observed.[5] In the case of Earth, the solar wind convects past the shock on a time scale of ~1 hour. Thus, even though the shock itself is a permanent feature, the actual lifetime of particle interactions with the shock is limited, resulting in the acceleration of particles to only modest energies.

Planets without magnetic fields (Mars, Venus) and comets also have bow shocks, which are produced by the interaction between the solar wind and the planetary or cometary ion populations. Cometary and planetary neutral atoms are ionized by sunlight and then picked up by the solar wind. The pickup ions mass load and slow the solar wind, eventually resulting in the formation of a bow shock. The pickup ions also generate intense wave activity. These waves on both sides of the shock appear to accelerate ions to higher energies, in a manner similar to that seen at other solar system shocks. The in situ plasma and energetic particle measurements acquired during the Giotto spacecraft's close encounter with comets Halley and Grigg-Skellerup and the ICE spacecraft's close encounter with comet Giacobini-Zinner clearly showed that ions are accelerated to energies of hundreds of keV in the cometary environment.

Termination Shock of the Solar Wind As the solar wind moves out of the inner heliosphere, the density of the plasma and the pressure that it exerts greatly exceed the pressure and density of the dilute, cold plasma in the local interstellar medium, and so the solar wind blows aside the interstellar plasma to form a cavity called the heliosphere (see Figure 3.5). The solar wind density decreases as the square of the distance from the Sun. Eventually, when the solar wind is roughly 10,000 times less dense than at Earth orbit, its pressure is so low that it cannot push aside the thin interstellar plasma, and it slows abruptly, creating a shock that marks the termination of the solar wind beyond which the shocked solar wind is diverted to form a heliosheath.[6] The solar wind plasma of the heliosheath is separated from the interstellar plasma by a discontinuity known as the heliopause, which is the heliosphere's outer boundary.

The actual size of the termination shock is unknown. The distance to the shock is estimated to be around 100 AU. In this distant region, the low-energy particle population most easily accelerated consists of interstellar neutral atoms that have penetrated deep into the heliosphere and come close enough to the Sun to lose one of their orbital electrons as a result of photoionization by solar ultraviolet radiation or through charge exchange with the solar wind. They are then picked up by the solar wind and transported outward, taking a year or more to reach the termination shock. Some of these ions then interact with the shock and move along its length for scores of astronomical units, eventually escaping from the shock with energies of up to tens of MeV/nucleon.[7]

The origin of the energetic particles from this so-far-unseen shock has been deduced from their highly unusual elemental composition, which led to their being called anomalous cosmic rays (ACRs). The main ACR source population consists of neutral atoms with a relatively high first ionization potential that allows them to enter the heliosphere as neutrals, become ionized, and then participate in the acceleration process when they reach the termination shock.

Virtually all of the heliospheric energetic particle populations observed outside magnetospheres and not associated with coronal active regions are associated with shocks. Where the shock is directly observable, the energetic particles are generally observed to have maximum intensity at the shock. The basic acceleration mechanism is essentially Fermi acceleration at either a quasi-parallel or a quasi-perpendicular shock wave. Coherent shock-drift acceleration also plays an important role.

Outstanding Questions About Shock Acceleration

- What are the seed populations for shock-accelerated particles and the injection mechanism?
- What are the sources and composition of pickup ions?
- What mechanism accelerates ions to energies greater than 1 GeV at the Sun?
- What causes the apparent plateau in the solar energetic particle intensity?
- What causes the universal spectra of solar energetic particle events at late times?
- What are the scalings that allow for detailed models developed, tested, and constrained in the heliosphere to be extended to remote cosmic environments such as supernova remnant shocks and the termination shocks of stellar winds?

COHERENT ELECTRIC FIELD ACCELERATION

Particles can be accelerated by means of large-scale, coherent (nonstochastic) electric fields. Conceptually, such acceleration is easy to understand when it occurs in the direction parallel to strong magnetic fields (as along magnetic field lines connected to Earth's auroral ionosphere) or in regions where the magnetic field strength is very weak (as perhaps near the center of the neutral sheet that separates the northern and southern magnetic lobes of Earth's magnetotail). When charged particles are placed in a vector electric field **E**, they experience a vector force (**F**), which has magnitude $F = qE$. A charged particle will accelerate or decelerate as long as this force is present, or as long as the particle is not diverted out of the acceleration region by the cumulative action of weak magnetic fields. For convenience this conceptually simple process of acceleration by parallel electric fields is referred to here as *direct electric field acceleration*.

Strong particle acceleration also occurs when coherent electric fields are applied in directions perpendicular to relatively strong magnetic fields, a process referred to here as *indirect electric field acceleration*. Indirect acceleration is either adiabatic or nonadiabatic. It is adiabatic when the temporal and spatial scales of the acceleration are large compared with the temporal and spatial scales associated with the gyrating and bouncing motions of the particles around and along the magnetic field lines (Figure 6.1). It is nonadiabatic when the temporal and spatial scales of the acceleration approach (or are smaller than) the scales of the gyration and bouncing motions of the particles.

In the adiabatic case, the acceleration of the particles is relatively easy to calculate. Within the complex geometry of a realistic magnetic field configuration (as found, for example, within a planetary

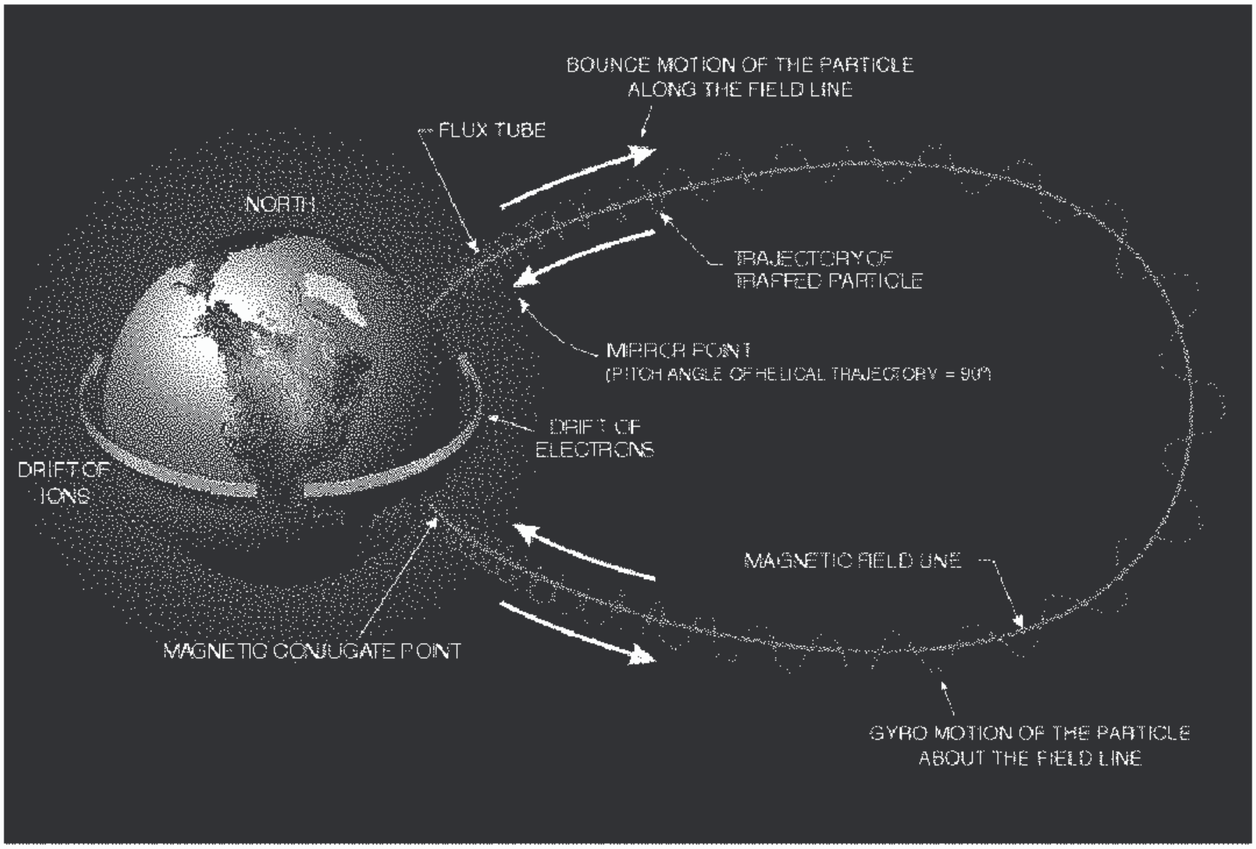


FIGURE 6.1 The three basic motions of a charged particle in a magnetic field are gryo, bounce, and drift. The "bouncing" occurs in closed magnetic geometries where particles bounce back and forth between regions along the magnetic field lines where the magnetic field strength is strong. Adapted from W.N. Spjeldvik and P.L. Rothwell, The radiation belts, *Handbook of Geophysics and the Space Environment*, A.S. Jursa, ed., Air Force Geophysics Laboratory, Air Force Systems Command, USAF, Hanscom AFB, Mass., 1985.

magnetosphere), there are "guiding center" drifts (with vector velocities $\mathbf{V}_D$) that arise strictly from the effect of the nonuniformities within the magnetic field. The acceleration of such particles engendered by a coherently applied electric field $\mathbf{E}$ is simply the vector dot product: $q\mathbf{E} \cdot \mathbf{V}_D$. The process results in acceleration parallel to the background magnetic field (betatron acceleration) and perpendicular to it (Fermi acceleration, a variant of the Fermi acceleration discussed in the context of shocks).

Nonadiabatic acceleration can produce even stronger particle energization than adiabatic acceleration can. If the electric field increases in a period of time that is commensurate with the time period associated with the bounce motions of the particles, or even faster and commensurate with the gyromotions of particles around the magnetic field, then the acceleration parallel or perpendicular to the background magnetic field can be much larger than that achieved by adiabatic betatron acceleration or Fermi acceleration. If the electric field turn-on is "instantaneous," then the perpendicular energization is similar to the so-called pickup energization that occurs when an atom is suddenly ionized in the presence of an electric field.

In space plasmas, the electric fields invoked in the preceding paragraphs can be produced either by a separation of positive and negative charges (as in the case of parallel auroral acceleration) or by the action of time-varying magnetic fields through Faraday's law of magnetic induction (as is the case when the magnetic tail of Earth's magnetosphere partially collapses and becomes more dipolar in configuration). The coherent electric field acceleration processes described above will occur in various manifestations in all dynamic magnetized plasma systems. As mentioned previously, direct electric field acceleration produces the energetic electrons that create auroras at Earth. Various combinations of direct and indirect electric field acceleration also operate at sites of magnetic reconnection such as in planetary magnetotails and other current-sheet structures, and in solar flares. In magnetospheres, coherent electric field acceleration helps create the pressure-bearing plasma that supports magnetotails against collapse, and helps create planetary radiation belts. The pickup energization mentioned above occurs, in the presence of the solar wind's motional electric field, to ions newly created by the ionization of neutral atoms from planetary or cometary atmospheres (cf. Chapter 4). The pickup process also accelerates those ions that enter the heliosphere originally as neutrals from the local interstellar medium. The subsections that follow expand on some of the most interesting examples of coherent electric field acceleration.

Radiation Belt Particles

Earth's Radiation Belts The acceleration of trapped particle populations in magnetospheres is facilitated by the residence time of particles in regions where energy can be extracted from time-varying fields. Particles that originate in the distant magnetic tail of a magnetosphere and the solar wind are often moved abruptly toward Earth as a result of bursty reconnection. Inductive electric fields produced by the abruptly changing magnetic fields in these flows accelerate electrons and ions prior to their entry into the radiation belts. These injected particles, associated with geomagnetic storms and substorms, constitute a source of the planetary radiation belts, but the quantitative contribution of this pre-acceleration process to the overall energization of radiation belt particles during storms is still not known. In Earth's magnetosphere, an extreme example of impulsive acceleration occurred March 24, 1991, when a high-speed interplanetary shock launched by a CME compressed the boundary of the magnetosphere well inside the orbit of geosynchronous spacecraft.[8] The event produced >10-MeV electron and proton radiation belts within a normally benign region called the electron slot (Figure 6.2). Trapping and energization to ~10 MeV occurred on a particle drift time scale of minutes as a result of the induction electric field launched by rapid magnetopause compression. The new >10-MeV electron and proton belts persisted for years.

The March 1991 event was unusual. The more usual process for generating or enhancing radiation belts is described here. Near solar maximum, large geomagnetic storms are often initiated by CME-driven interplanetary shocks that compress the magnetosphere, causing what is called a storm sudden commencement. After an initial compression by the shock and during an extended interval when the southward orientation of the interplanetary magnetic field results in the efficient transfer of energy from the solar wind to the magnetosphere, magnetospheric plasma is transported radially inward from the plasma sheet (Figure 6.3) to a region of stronger magnetic field, where the radial gradient in the magnetic field causes a westward drift of energetic (tens to hundreds of keV) ions and an eastward drift of electrons ($\mathbf{V}_D$ in the preceding section) producing the so-called ring current (Figure 6.3). When such drifting particles experience a variation in the magnetic field at a frequency comparable to that of its drift period, they will diffuse

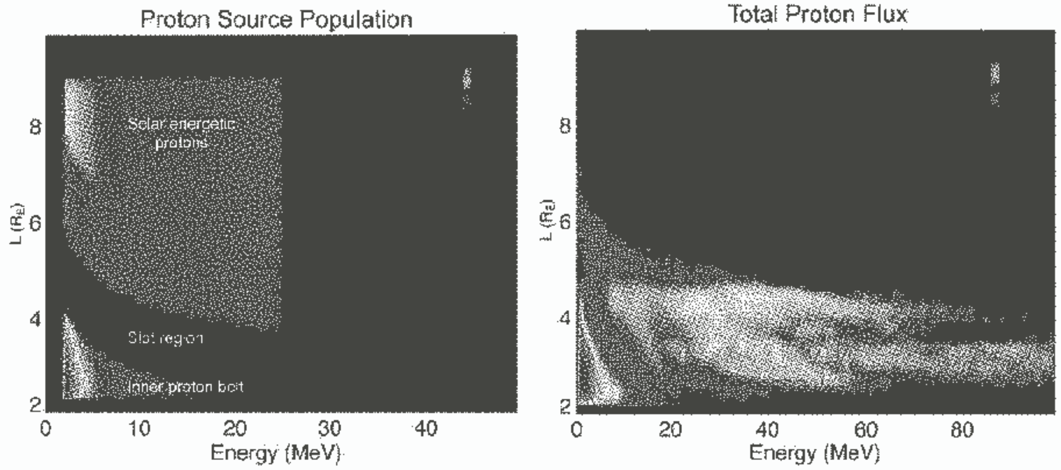


FIGURE 6.2 Computer simulation of the formation of a new proton belt during the magnetic storm of March 24, 1991. The panel on the left shows the seed population of solar energetic protons, while the panel on the right shows the new proton belt formed in the slot region as a result of the interaction of the CME-driven interplanetary shock with Earth's magnetosphere. Image courtesy of M.K. Hudson (Dartmouth University). Modified and reprinted, with permission, from M.K. Hudson, Simulations of radiation belt formations during storm sudden commencements, *Journal of Geophysical Research* 102(A7), 14087-14102, 1997. Copyright 1997, American Geophysical Union.

---

radially from one drift shell to another. Recently, ultralow-frequency waves have been shown to play a key role in the inward diffusion of particles. When the particles diffuse planetward, where the magnetic field increases in strength, their energies increase.

Coherent electric field acceleration cannot explain all aspects of radiation belt enhancement. Satellite measurements of particle velocity distributions suggest that additional heating due to waves with frequencies comparable to the electron gyrofrequency is taking place.

Other Planetary Magnetospheres Planetary probes have identified trapped relativistic electrons and energetic ions in the magnetospheres of Jupiter, Saturn, Uranus, and Neptune. Processes both similar to and distinct from those occurring at Earth energize these particle populations. Jupiter's radiation belts are the most intense in the solar system, as much as a factor of 1000 more intense than Earth's. The combination of that intensity and the strength of Jupiter's magnetic field (with a magnetic moment 20,000 times that of Earth's) results in the emission of observable synchrotron radiation in the decimetric wavelength range, a unique characteristic of Jupiter's magnetosphere (Figure 6.4). Jovian radiation belts derive their high energies (1000 MeV) and flux levels from the planet's rotational kinetic energy and are largely shielded from the buffeting by the interplanetary environment that plays such a critical role in generating the terrestrial radiation belts. Data acquired during the 8-year orbital tour of the Jupiter system by the Galileo spacecraft has revolutionized our knowledge of many aspects of the behavior of the jovian magnetosphere. However, how Jupiter, utilizing steady rotational energies rather than the dynamic solar wind, generates such powerful and energetic radiation regions remains obscure. Detailed information about the radiation belt environment of another giant outer planet will become available when the Cassini orbiter begins its

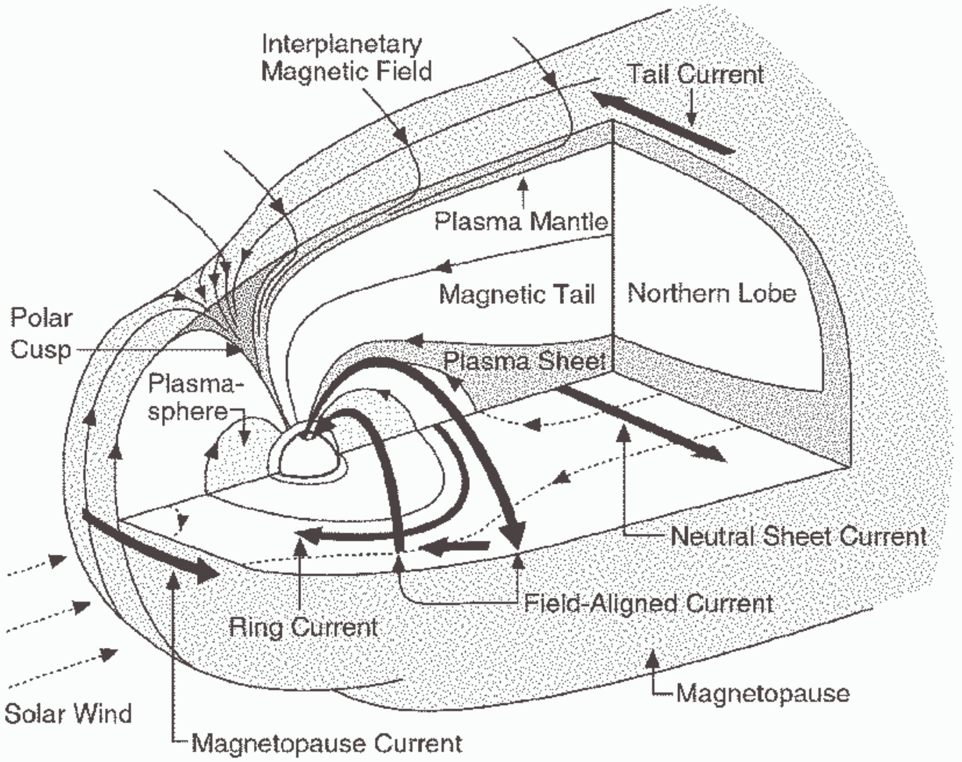


FIGURE 6.3 Schematic of the magnetosphere showing the principal plasma populations and current systems. Courtesy of C.T. Russell, University of California, Los Angeles.

tour of the Saturn system in mid-2004. The data provided by Cassini on the energetic particle populations in Saturn's inner magnetosphere will be an invaluable contribution to comparative studies of the acceleration, transport, and loss of radiation belt particles in different planetary environments.

Solar Flares

Hard x-ray/gamma-ray continuum and gamma-ray line observations show that solar flares, as well as fast CMEs, can accelerate ions up to tens of GeV and electrons to hundreds of MeV. Flares release up to 10^{32} to 10^{33} ergs in 10^2 to 10^3 s, with the accelerated 10- to 100-keV electrons (and probably $\gtrsim 1$ MeV/nucleon ions) containing a significant fraction, ~10 to 50 percent, of this energy. How the Sun releases this energy, presumably stored in the magnetic fields of the corona, and how it rapidly accelerates electrons and ions with such high efficiency, and to such high energies, are currently not known. Hard x-ray spectra obtained with high spectral resolution show a break at ~20 to 100 keV, suggesting that the accelerated electrons have a sharp feature in that energy range.[9] Similar features in electron spectra observed in Earth's auroral zone are the result of acceleration by a quasi-stationary (DC) electric field parallel to **B**, with the peak energy corresponding to the total potential drop. Coherent, and perhaps direct, electric field accelera-

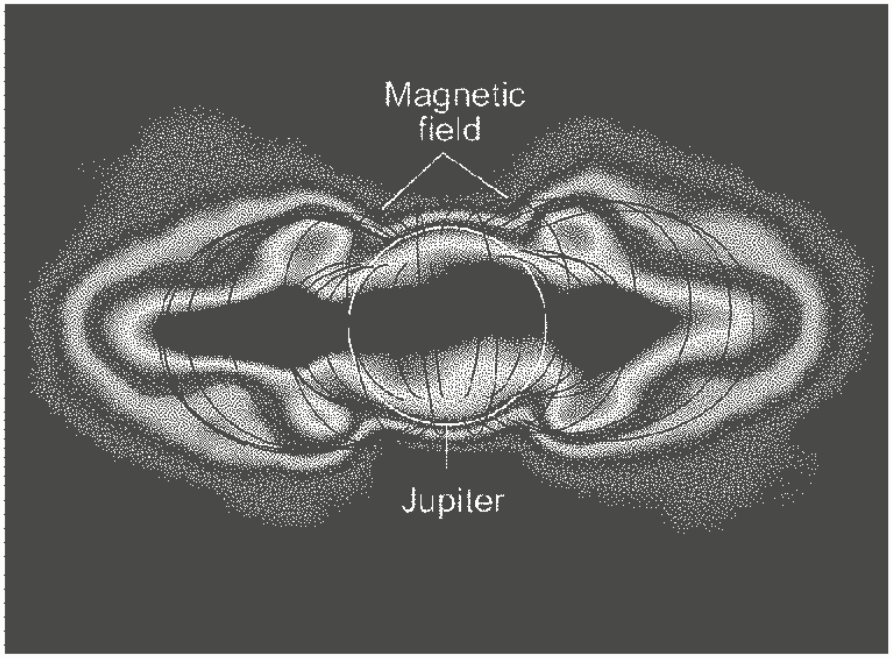


FIGURE 6.4 Energetic electrons accelerated to velocities near the speed of light are responsible for the synchrotron emission from Jupiter's powerful radiation belts. The false color indicates the intensity of the emission, with red being the most intense. Courtesy of Imke de Pater (University of California, Berkeley), NRAO/VLA, and Sky Publishing Co.

tion is thus believed to play a role in particle acceleration associated with solar flares, although other processes (shock acceleration and stochastic acceleration by magnetohydrodynamic (MHD) turbulence) are also thought to be involved.

The Auroral Magnetosphere

Electric fields parallel to the background magnetic field have long been thought to play a role in producing Earth's aurora. A field-aligned potential drop was first proposed in the mid-1970s by David Evans to explain the monoenergetic electron beams observed in association with auroral arcs.[10] Evidence for the existence of such structures has been provided by double probe measurements, chemical release experiments, and particle data.

Prior to these measurements, debate focused on the implied violation of the frozen-in magnetic field condition of ideal magnetohydrodynamics, which requires that in a collisionless plasma any electric field **E** must be perpendicular to **B**. It was understood that at the lowest altitudes in the ionosphere, this condition is violated by collisions with neutrals, which permit ions to flow parallel to **E** and carry a current, while electrons flow perpendicular and in the **E** × **B** direction. The new understanding is that the ideal MHD conditions are violated well above the ionosphere where the plasmas are collisionless. In the ionosphere, the transverse currents allow closure of magnetic field-aligned currents imposed from the magnetosphere.

It has been shown recently that large-amplitude, coherent MHD (Alfvén) waves in the boundary layer between the central plasma sheet and tail lobes (see Figure 6.3), mapping to the auroral ionosphere, carry a downward wave energy flux that is consistent with the energy associated with the auroral emissions. Also, on large transverse scales compared to the ion gyroradius, perpendicular electric fields map along auroral field lines, while at smaller transverse scales a parallel potential drop can be inferred. The distribution of that potential drop along **B** continues to be debated, as does the quasi-static versus electromagnetic nature of the potential drop. Localized structures appear to provide an anomalous resistivity that modifies the reflection properties of Alfvén waves, carrying a parallel electric field component on transverse scales comparable to those of auroral arcs.

Outstanding Questions About Coherent Electric Field Acceleration

- What is the relative contribution to radiation belt acceleration and loss of magnetic-moment-conserving ultralow-frequency (MHz range) waves and other adiabatic transport processes versus nonconserving influences such as very low frequency (up to kHz) waves?
- How probable are extreme radiation belt flux enhancements such as that on March 24, 1991, which produced new MeV electron and trapped solar proton belts on a drift time scale of minutes?
- Why did outer zone electron fluxes essentially disappear for 2 months following the May 11, 1999, period in which solar wind density dropped to less than 0.1 cm^{-1}?
- How does Jupiter generate its incomparably powerful radiation belt in the absence of solar wind buffeting effects?
- What is the distribution of parallel electric field along **B** within the auroral acceleration region and how is it maintained?

STOCHASTIC PARTICLE ACCELERATION

In his original model for the acceleration of cosmic rays in the interstellar medium, Fermi suggested that the random movement of magnetic scattering centers or clouds could further energize fast-moving particles since they would experience more head-on (energy-gaining) than overtaking (energy-losing) collisions. In the interplanetary or interstellar medium, particles experiencing scattering by a random ensemble of waves or turbulence can effectively experience head-on and overtaking collisions. In this case, the process is a little more subtle. Consider for simplicity Alfvén waves only. One can introduce a frame of reference in which the motional electric field is transformed away so that a particle experiences pitch-angle scattering by the Alfvén waves and experiences diffusion in momentum space. Back in the laboratory frame of reference, researchers recover the motional electric field and find that the particles will have gained or lost energy in a somewhat random "stochastic" sense. The physical content of stochastic acceleration amounts to a particle being either accelerated or decelerated in randomly oriented electric field perturbations associated with the ensemble of Alfvén (or other) waves or turbulence.

One can show that the rate of energy gain is proportional to the square of the ratio of Alfvén speed to particle speed (a small number for energetic particles) and hence is often referred to as second-order Fermi acceleration. By contrast, the presence of a shock wave ensures that all particle "collisions" are effectively head-on. The energy gain in this case is proportional to the ratio of the shock speed to the particle speed rather than the square of the Alfvén to particle speed. Consequently, shock acceleration is referred to as first-order Fermi acceleration and is generally much more efficient than stochastic acceleration.

The most detailed studies of stochastic acceleration have been based on assuming either Alfvénic or slab turbulence or low-frequency MHD waves. The relatively simple relationships between velocity and magnetic field fluctuations allow tractable forms of the momentum diffusion coefficient to be derived and included in models of particle transport in the solar wind, for example. Such models have addressed the origin and transport of solar energetic particles and interstellar pickup ions. However, the characteristics of energetic particles observed in situ, while admitting a partial explanation in terms of stochastic acceleration, continue to defy simple theoretical explanation. For example, accelerated pickup He^+ spectra as observed at 1 AU reveal the existence of (1) a knee connecting the pickup ion "core" with the accelerated "tail" and (2) an accelerated tail with a rather flat slope. The accelerated tail closely resembles that which might be expected from diffusive shock acceleration. However, the observations are integrated over relatively long periods when the solar wind was especially quiet and free of interplanetary shocks. Simple stochastic acceleration models based on slab turbulence have considerable difficulty explaining such observations.

Alternative approaches based on a much more sophisticated description of interplanetary or interstellar turbulence are now under consideration for stochastic acceleration models. Observations, theory, and simulations all suggest that MHD turbulence in the solar wind is quasi-two-dimensional, which is superimposed on a large-scale interplanetary magnetic field. One interesting feature of two-dimensional turbulence is that random convective motions of MHD vortices lead to nonlinear interactions of neighboring magnetic islands. Turbulent reconnection between neighboring magnetic islands of opposite magnetic polarity creates turbulent electric fields. Very strong turbulent intermediate-scale electric fields have been observed in the solar wind at low helio-latitudes, and simulations of particles in a two-dimensional turbulence field demonstrate that turbulent electric fields can efficiently accelerate charged particles. Turbulent electric fields associated with two-dimensional turbulence could explain the acceleration of pickup and solar wind ions in the quiet low-latitude solar wind, i.e., in the absence of nearby shocks. Clearly, to understand stochastic acceleration in realistic physical environments requires an intimate understanding of local turbulence.

Outstanding Questions About Stochastic Particle Acceleration

- What is the origin of the power law-like accelerated ion tail in the quiet solar wind?
- Can a stochastic acceleration process provide the seed particles for diffusive shock acceleration?
- What is the effect of different turbulence characteristics on stochastic particle acceleration?
- In what plasma regimes is stochastic particle acceleration effective compared to alternative acceleration processes?

SUMMARY

In situ access to energetic particle acceleration mechanisms in the solar system provides a unique opportunity to make direct measurements of processes that can be scaled up to astrophysical counterparts.

The rapid acceleration of energetic particles in solar flares and substorms is still not well understood. Undoubtedly, reconnection processes discussed in Chapter 2 lead to explosive energy conversion, discussed in Chapter 5, and a substantial fraction of that energy is carried away by energetic particles.

Slower acceleration by interplanetary shocks, typically initiated by CMEs inside 1 AU, is better understood at the level of local shock acceleration, but the current level of understanding does not yet provide a macroscopic model for solar energetic particle events, nor explain quantitatively the observed plateau in maximum flux. We do not yet have a quantitative predictive capability for solar energetic particle fluxes in the magnetosphere, nor of their trapping lifetimes, which requires a better understanding of trapping and loss processes.

The relative importance of various radial transport processes and of localized acceleration in radiation belt enhancements is not well understood. An improved understanding of these processes is needed so that researchers can better characterize the harsh and highly time-variable radiation environment of the inner magnetosphere, where many important spacecraft operate, and so that potentially hazardous conditions can eventually be predicted. We do not know, for example, how probable an extreme event such as the March 24, 1991, enhancement is, which produced fluxes still measurable a decade later in a generally benign region of the radiation belts (the so-called slot region for its usual absence of flux).

Direct acceleration by parallel electric fields is important at the Sun and in auroral acceleration regions. While much progress has been made in understanding microphysical processes of parallel electric field acceleration, there is still much uncharted territory. So, for example, it was learned only recently how structured and time-variable the auroral return current region is, and particle acceleration theory has not yet caught up with the recent observations from the FAST satellite. How such fields are modulated on the macroscale is not yet well understood.

Finally, it is important to note the universality of the particle acceleration mechanisms described here. Particle acceleration provides one of the most outstanding examples of cross-fertilization between the space physics and astrophysics communities. Mechanisms developed in one community have almost inevitably migrated to the other community, the quintessential example being diffusive shock acceleration. The virtue of the heliosphere is that it allows for detailed in situ investigation of acceleration processes, whereas the universe allows us to consider (and possibly test remotely on the basis of photons—synchrotron radiation emitted by energized electrons, for example) much more extreme environments and to explore the possible scalings from heliosphere to galaxy. In a general sense, the development of such scalings remains as one of the outstanding questions related to particle acceleration and one that emphasizes the universality of particle acceleration throughout the cosmos.

NOTES

1. Enrico Fermi originally envisioned the acceleration of galactic cosmic rays by sequential "bounces" off interstellar magnetic clouds in relative motion. Galactic cosmic rays are generally thought to be accelerated by the Fermi mechanism at shocks associated with supernova remnants. Cf. E. Fermi, On the origin of the cosmic radiation, *Physical Review* 75, 1169-1174, 1949. The modern theory of diffusive shock acceleration was developed in the late 1970s (see the review by W.I. Axford, Particle acceleration on galactic scales, pp. 45-56 in *Particle Acceleration in Cosmic Plasmas*, AIP Conference Proceedings 264, G.P. Zank and T.K. Gaisser, eds., American Institute of Physics, New York, 1992.

2. T.P. Armstrong, M.E. Pesses, and R.B. Decker, Shock drift acceleration, pp. 271-285 in *Collisionless Shocks: Reviews of Current Research*, Geophysical Monograph 35, B.T. Tsurutani and R.G. Stone, eds., American Geophysical Union, Washington, D.C., 1985.

3. J.T. Gosling, Coronal mass ejections: An overview, pp. 9-16 in *Coronal Mass Ejections*, Geophysical Monograph 99, N. Crooker, J.A. Joselyn, and J. Feynman, eds., American Geophysical Union, Washington, D.C., 1997.

4. M.I. Desai et al., Particle acceleration at corotating interaction regions in the three-dimensional heliosphere, *Journal of Geophysical Research* 103, 2003, 1998; J. Giacolone, J.R. Jokipii, and J. Kóta, Particle acceleration in solar wind compression regions, *Astrophysical Journal* 573, 845-850, 2002.

5. C.F. Kennel, Collisionless shocks and upstream waves and particles: Introductory remarks, *Journal of Geophysical Research* 81, 4325, 1981.

6. J.R. Jokipii and F.B. McDonald, Quest for the limits of the heliosphere, *Scientific American* 272, 58, 1995.

7. B. Klecker et al., Anomalous cosmic rays, *Space Science Review* 83, 259, 1998.

8. Before and after snapshots of the shock arrival from a global magnetohydrodynamic simulation of this event appear in Plate 4 of the NRC report *Radiation and the International Space Station: Recommendations to Reduce Risk*, National Academy Press, Washington, D.C., 2000.

9. R.P. Lin et al., RHESSI observations of particle acceleration and energy release in an intense solar gamma-ray line flare, Astrophysical Journal Letters 595(2), L69-L76, 2003.

10. D.S. Evans, Precipitating electron fluxes formed by a magnetic field aligned potential difference, *Journal of Geophysical Research* 79, 2853-2858, 1974.

7

Concluding Thoughts

Our solar system, and stellar systems in general, are rich in the dynamical behaviors of plasma, gas, and dust organized and affected by magnetic fields. These dynamical processes are ubiquitous to highly evolved stellar systems, such as our own, but also play important roles in their formation and evolution. Stellar systems are born out of clumpy, rotating, primordial nebulas of gas and dust. Gravitational contraction, sometimes aided by shock waves (possibly from supernovas), passage through dense material, and other disruptions, forms condensation centers that eventually become stars, planets, and small bodies. Magnetic fields moderate early-phase contractions and may also play vital roles in generating jets and shedding angular momentum, allowing further contraction. The densest of the condensation centers become protostars surrounded by accretion disks. Dynamo action occurs within the protostars as the heat of contraction ionizes their outer gaseous layers, resulting in stellar winds. In similar fashion, rotating solid and gaseous planets form, and many of these also support dynamo action, producing magnetic fields. Ultraviolet and x-ray photons from the central stars partially ionize the upper atmospheres of the planets as well as any interstellar neutral atoms that traverse the systems. Viewed as a whole, the resulting plasma environments are called asterospheres, or in the Sun's case, the heliosphere. In its present manifestation, the heliosphere—the local cosmos—is a fascinating corner of the universe, challenging our best scientific efforts to understand its diverse machinations. It must be appreciated at the same time that our local cosmos is a laboratory for investigating the complex dynamics of active plasmas and fields that occur throughout the universe from the smallest ionospheric scales to galactic scales. Close inspection and direct samplings within the heliosphere are essential parts of the investigations that cannot be carried out by a priori theoretical efforts alone.

This report summarizes much of what is known about the plasma physics of the local cosmos and lists many of the outstanding questions that will be driving the field for the near future. The discussions are organized around five broad themes, specifically (1) the creation and annihilation of magnetic fields, (2) the formation of structures and transients, (3) plasma interactions, (4) explosive energy conversion, and (5) energetic particle acceleration. These phenomena have been identified, and questions posed, in terms of specific observables either on the Sun or in various parts of the heliosphere and planetary systems. The proposed solutions and experiments are also designed for the parameter ranges found in the local cosmos. Nonetheless, in every case the solutions are of universal applicability, and a challenge for solar and space physics in the future is to extend the models and theories that have been validated in the local cosmos to nonlocal astrophysical plasmas. It is obvious, for example, that the theories of magnetic dynamos and magnetic reconnection must be applicable to the generation of magnetic fields and to the explosive conversion of magnetic energy to heat and particle kinetic energy in every corner of the universe. Likewise, the acceleration of charged particles to high energies by shock waves and by both parallel and perpendicular electric fields is certainly of universal importance. Obstacles to plasma flows exist everywhere, and the interaction between two magnetized plasma populations can produce magnetic stresses that are relieved with cross-scale coupling processes that occur either gradually or explosively depending upon the capabilities of each plasma region.

As the discipline of solar and space physics has matured, the focus has become less on places to explore than on fundamental processes to investigate and understand. Understanding requires that the processes be investigated in diverse plasma environments. Important to this investigation are space missions to the magnetospheres of other planets as well as that of Earth, missions to sample the properties of the heliosphere, missions to observe the astonishing fine structure of the active Sun, and multispacecraft missions to sample the structure of magnetopauses, shock transitions, and the magnetic reconnection processes. The results must be fed into the latest theoretical models, and it is these models that provide the links to other parts of the cosmos. Thus, a plasma theory and modeling program that cuts across the disciplines of solar physics, space physics, and astrophysics is an important part of any efforts to understand the plasma physics of the cosmos.

Appendixes

A

Statement of Task

Background Space is filled with magnetized plasma. In its natural cosmic setting in the solar system, magnetized plasma is now known to display a set of characteristic structures and processes that in turn have characteristic modes of behavior. These structures and processes occur with vast ranges of size, duration, and energy that are self-organized into distinct classes of phenomena.

The Sun is a major source of energy and magnetized plasma in the solar system. As such, it has important connections to astrophysics and to the space environment near Earth. Four decades of space exploration have measured and recorded plasma behavior near Earth and at many solar system objects, including 7 planets, 6 satellites, 2 comets, and 2 asteroids. Other spacecraft have measured the solar wind from heliocentric pole to pole and from the orbit of Mercury to the outermost recesses of the heliosphere, recording sundry indigenous structures and processes. In addition, space-borne telescopes have revealed the Sun's features and movements at ever more wavelengths and higher resolutions.

NASA is currently planning an ambitious program of future missions that promise to further reveal how magnetized plasmas are organized in space and how they behave. With a rich data legacy and a promising measurement future, there now exists the opportunity to foster a new disciplinary thrust in space and solar physics, one that will emphasize that the locally occurring (solar system) structures and processes also have astrophysical counterparts and are, in fact, characteristic of cosmic plasma behavior. The committee refers to this evolving branch of space and solar physics as "solar connections."

Plan The committee will undertake a study with the following objectives:

- *Explicate the content of solar connections.* The CSSP will outline the underlying scientific basis for contemporary solar system plasma physics, identify major outstanding scientific questions, and define the interface or links to studies of astrophysical plasmas at one extreme and the NASA Sun-Earth Connection/Living With A Star programs on the other.
- *Assess the field's current data, theory, and computational resources as they pertain to solar connections.*
- *Recommend measures, including but not restricted to missions, to further develop the field.*

To acquire background information for the report, the committee will form study groups organized around 5 themes. These themes, which may evolve over the course of the study, comprise a convenient, but not unique, scientific framework around which to structure an assessment of the key physical processes of interest. The themes are:

1. Creation and Annihilation of Magnetic Fields
2. Spontaneous Generation of Structures and Transients
3. Magnetic Coupling
4. Explosive Energy Conversion
5. Generation of Penetrating Radiation

Each study group will consist of 2-3 committee members and several experts from the science community. Each group will compile a comprehensive set of examples, structures, and processes that belong to their theme. They will then define the field's data and theoretical/computational requirements and assess critically the field's scientific potential. Finally, the study groups will suggest directions likely to produce the greatest advances, and note what missions, planned or as yet unplanned, are needed to promote the advancement.

In generating its report, the full committee will draw upon the findings and recommendations of the study groups. In addition to defining the content of solar connections, the report will evaluate planned NASA missions in terms of their relevance to solar connections. Where necessary, the committee will also recommend additional missions or priorities. In particular, the committee will critically evaluate the SEC plan and identify areas where enhanced theoretical-computational emphasis is needed to properly support the solar-connections effort.

B

Study Groups

Creation and Annihilation of Magnetic Fields
James L. Burch, Southwest Research Institute, *Study Organizer*
James F. Drake, University of Maryland
Michael Hesse, NASA Goddard Space Flight Center
Weijia Kuang, NASA Goddard Space Flight Center
Eugene N. Parker, University of Chicago, Professor Emeritus

Formation of Structures and Transients
Judith T. Karpen, Naval Research Laboratory, *Study Organizer*
Charles W. Carlson, University of California, Berkeley
Robert L. Carovillano, Boston College
Robert E. Ergun, University of Colorado, Boulder
Thomas W. Hill, Rice University
Jack D. Scudder, University of Maryland

Plasma Interactions
Robert W. Schunk, Utah State University, *Study Organizer*
Tom Chang, Massachusetts Institute of Technology
Paul A. Cloutier, Rice University
Thomas L. Holzer, U.S. Geological Survey
David G. Sibeck, Johns Hopkins University
Richard A. Wolf, Rice University

NOTE: Affiliations listed were current at the time of the meetings of the study groups.

Explosive Energy Conversion
George Siscoe, Boston College, *Study Organizer*
Spiro K. Antiochos, Naval Research Laboratory
Tom Chang, Massachusetts Institute of Technology
Ferdinand V. Coroniti, University of California, Los Angeles
Terry G. Forbes, University of New Hampshire
Joachim Raeder, University of California, Los Angeles
Vytenis M. Vasyliunas, Max-Planck-Institut fuer Aeronomie
Richard A. Wolf, Rice University

Energetic Particle Acceleration
Mary Hudson, Dartmouth College, *Study Organizer*
Steven Kahler, Air Force Research Laboratory
Martin A. Lee, University of New Hampshire
Robert P. Lin, University of California, Berkeley
Glenn M. Mason, University of Maryland
Barry H. Mauk, Johns Hopkins University
Frank B. McDonald, University of Maryland

C

Acronyms and Abbreviations

ACR	anomalous cosmic ray
AU	astronomical unit (150,000,000 km)
AURA	Associated Universities for Research in Astronomy
CIR	corotating interaction region
CME	coronal mass ejection
DNL	distant neutral line
ESA	European Space Agency
FAST	Fast Auroral Snapshot Explorer (satellite)
GCR	galactic cosmic ray
ICE	International Cometary Explorer (spacecraft)
IMAGE	Imager for Magnetopause-to-Aurora Global Exploration (satellite)
IMF	interplanetary magnetic field
LISM	local interstellar medium
MHD	magnetohydrodynamic
NASA	National Aeronautics and Space Administration
NENL	near-Earth neutral-line (model)
NRAO	National Radio Astronomy Observatory
R_m	(nondimensional) magnetic Reynolds number
SOHO	Solar and Heliospheric Observatory
STScI	Space Telescope Science Institute
TRACE	Transition Region and Coronal Explorer
VLA	Very Large Array