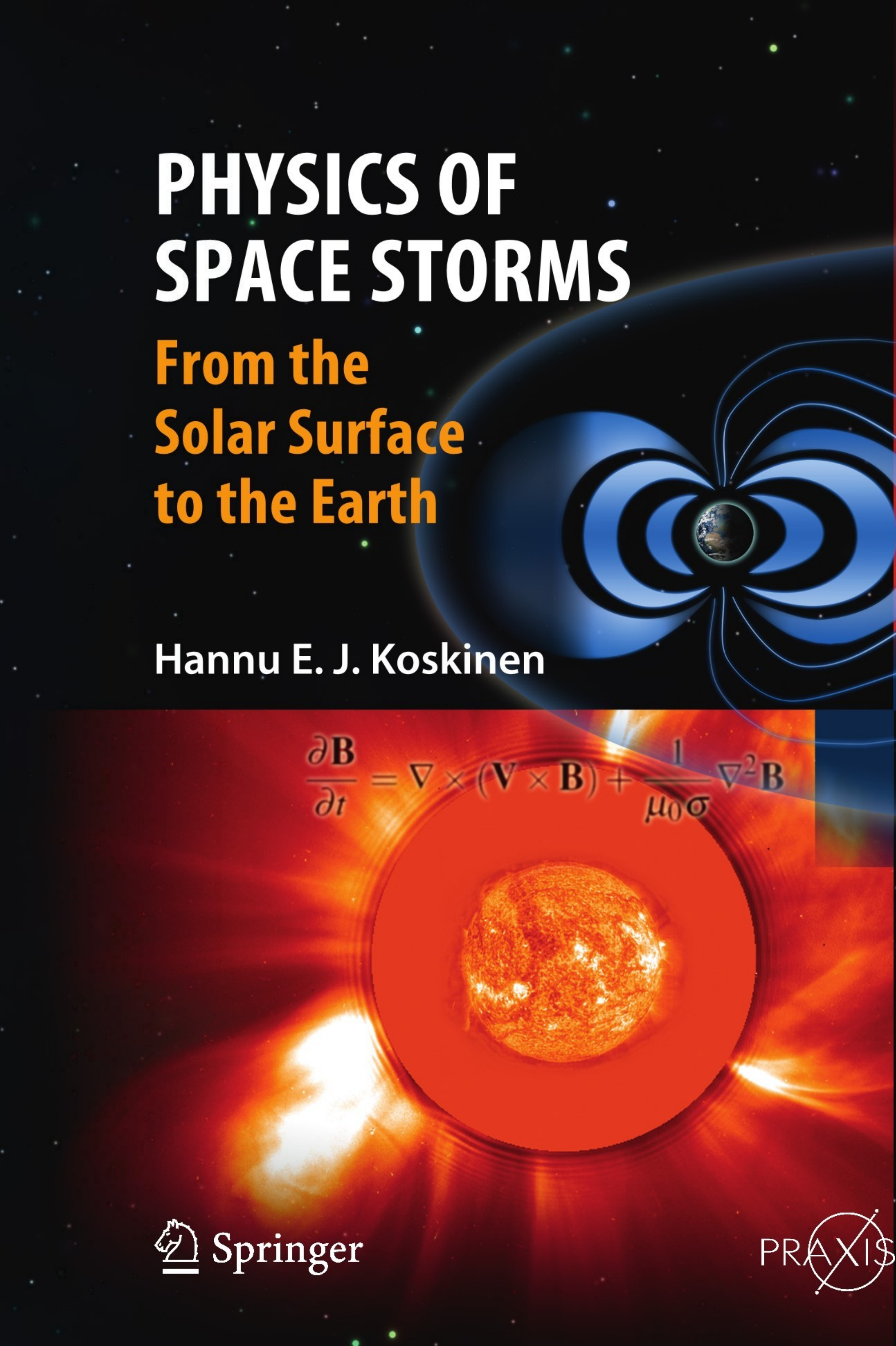

PHYSICS OF SPACE STORMS

From the Solar Surface to the Earth

Hannu E. J. Koskinen

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{V} \times \mathbf{B}) + \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B}$$

Springer

PRAXIS

Physics of Space Storms

From the Solar Surface to the Earth

Hannu E. J. Koskinen

Physics of Space Storms

From the Solar Surface to the Earth

Springer

Published in association with
Praxis Publishing
Chichester, UK

PRAXIS

Professor Hannu E. J. Koskinen
University of Helsinki and Finnish Meteorological Institute
Helsinki
Finland

SPRINGER–PRAXIS BOOKS IN ENVIRONMENTAL SCIENCES
SUBJECT *ADVISORY EDITOR*: John Mason, M.B.E., B.Sc., M.Sc., Ph.D.

ISBN 978-3-642-00310-3 e-ISBN 978-3-642-00319-6
DOI 10.1007/978-3-642-00319-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010934386

Cover design: Jim Wilkie
Project copy editor: Mike Shardlow
Author-generated LaTex, processed by EDV-Beratung Herweg, Germany

Printed on acid-free paper

Springer is part of Springer Science + Business Media (www.springer.com)

Contents

Preface

Space weather can be defined as a subtopic of solar–terrestrial physics, which deals with the spatially and temporally variable conditions in the Sun, solar wind, magnetosphere, and ionosphere that may disturb or damage technological systems in space and on the ground and endanger human health. *Space storms* are the strongest and most harmful appearances of space weather.

During the 1990s space weather grew to a prominent, if not the dominant, sector within solar–terrestrial physics. Also a significant fraction of basic space plasma physics research became motivated by its potential to contribute to useful space weather applications including more accurate forecasts. A key reason for the evolution of space weather activities is the growing understanding that a great number of systems in space, human beings included, and on the ground are vulnerable to severe space weather conditions. In fact, due to miniaturization and increasing complexity many technological systems are becoming more sensitive to the radiation environment than before. At the same time modern society is getting increasingly dependent on space infrastructure. In future the human presence in space, including space tourism, is expected to become more prominent. Some day we most likely will return to the Moon and, perhaps, initiate manned missions to Mars. On the ground the effects of space storms, such as saturation of transformers in electric power transmission networks or perturbations in telecommunication and global positioning systems, may be easier to handle, but this requires that the underlying physics be understood much better than today.

The developers of space weather services have done their best to follow the needs, sometimes real, sometimes imagined, of potential users of space weather applications. There is growing activity to produce tools for modeling and forecasting space weather conditions based on a limited set of observations, for specification of environmental conditions during storms, and for after-the-fact analysis of anomalous behavior of technological systems and hazards caused by severe space weather. Unfortunately, this activity is often based on insufficient knowledge of the underlying physical systems, sometimes even at the cost of basic research aiming at increasing this knowledge. This development is not always healthy in the long-term perspective. Furthermore, it is not enough just to solve the acute problems: the knowledge being gained today also needs to be maintained tomorrow.

While a large number of research articles and review papers on space storms have been published over the last several years, there is no comprehensive systematic textbook approach to the relevant physics of the entire chain of phenomena from the surface of the Sun to the Earth. The goal of the present monograph is to fill this gap. The text is aimed at doctoral students and post-doctoral researchers in space physics who are familiar with elementary plasma physics and possess a good command of classical physics. The topics reach from the storms in the solar atmosphere through the solar wind, magnetosphere, and ionosphere to the production of the storm-related geoelectric field on the ground. In the selection of material, preference has as much as possible been given to analytical and quantitative presentation over handwaving, while keeping the volume of the book reasonable.

Of course, several good plasma physics textbooks are available, which are useful in the education of space physicists, e.g., the rewritten classic of Boyd and Sanderson [2003], the little more challenging Sturrock [1994], or the recent volumes written by Gurnett and Bhattacharjee [2004] and Bellan [2006]. However, these books are written for very wide audiences from laboratory and fusion communities to space plasma physicists. Consequently, many important issues in the physics of tenuous space plasmas have had to be dealt with in a brief and cursory manner. For astrophysicists interested in the most abundant form of conventional matter in the universe the book by Kulsrud [2005] is strongly recommended, although quite demanding reading. There are also several textbooks with a clear focus on fundamental space plasma physics [e.g., Baumjohann and Treumann, 1996; Treumann and Baumjohann, 1996; Parks, 2003], but their approach too is more general than the thematically focused topic of the present volume. The multi-authored textbook edited by Kivelson and Russell [1995] covers large parts of the physical environment of this book. However, it does not go very deeply into the plasma physics and suffers to some extent from the different styles of the individually written chapters.

The rapid growth of space weather activities has led to a large number of compilation works of highly variable quality. An inherent problem of multi-authored collections is that each article is relatively short but at the same time written in a complete article style from introduction to conclusions and often with individual reference lists. Thus the books easily become thick but none of the articles can penetrate the basic physical principles. Some of the most useful collections in the present context are those edited by Crooker et al [1997], Tsurutani et al [1997], Daglis [2001], Song et al [2001], Scherer et al [2005], Baker et al [2007], Bothmer and Daglis [2007], and Lilensten et al [2008]. These books contain many excellent articles and provide students with a large body of study material with up-to-date observational data. However, these volumes rather complement than compete with this self-contained monograph.

This book can be interpreted to consist of three parts. The long Chapter 1 forms the first part. It contains a phenomenological introduction to the scene, from the Sun to the Earth, where space weather plays are performed. A reader familiar with basic physics of the Sun, solar wind, magnetosphere and ionosphere can jump over this chapter and only return to it when there is a need to check definitions or concepts introduced there.

The second part of the book consists of several chapters on fundamental space plasma physics. While this part is written in a self-consistent way, it is aimed at readers who already have been exposed to basic plasma physics. Chapter 2 briefly introduces the fundamental concepts and tools of plasma physics inherited from both electrodynamics and statistical physics. Chapter 3 reviews the classical guiding center approach to single particle motion and adiabatic invariants, including motion in the dipole field, near a current sheet, and in a time-dependent electric field.

Common problems to all plasma physics texts are in what order the microscopic and macroscopic pictures should be introduced and at what stage the waves and instabilities be discussed. The strategy in the present volume is to start with the wave concepts in the cold plasma approximation in Chapter 4. The chapter includes a discussion of radio wave propagation in the ionosphere as an example of dealing with wave propagation in inhomogeneous media in the WKB approximation, which is a powerful theoretical tool in problems where the wavelength is short as compared to the gradient scale lengths of the background parameters. Chapter 5 is a standard discussion of the Vlasov theory starting from Landau's solution and extending to the wave modes in uniformly magnetized plasma. Only after these is magnetohydrodynamics (MHD) treated in Chapter 6. Here more emphasis is placed on the field-aligned currents (i.e., force-free fields) than in many other plasma physics texts because they are of such great importance in the solar atmosphere, solar wind, and magnetosphere and in magnetosphere–ionosphere coupling. The chapter is concluded with a brief peek beyond the MHD approximation, including a quasi-neutral hybrid approach and the introduction of kinetic Alfvén waves.

Space plasma instabilities are the topic of Chapter 7. In whatever way you approach this complex, you end up being incomplete if you wish to keep the discussion within reasonable limits and focused. Here the approach is to introduce the basic ideas, such as the free-energy sources and stability criteria, behind several of the most important instabilities studied in the context of space storms, but most of the long and tedious derivations of the equations have been omitted. The reader interested in the details is recommended to consult more advanced textbooks in plasma theory and relevant research articles. Another choice motivated by the theme of this book is to discuss the magnetic reconnection and the tearing modes separately from other instabilities in a dedicated Chapter 8. Whatever the microphysical mechanisms associated with reconnection are, the understanding of its basic characteristics is an essential part of literacy in space physics, regardless of whether one is interested in solar flares, coronal mass ejections, solar wind interaction with the magnetosphere, or the substorms therein. Unlike other textbooks, the concept of dynamo is introduced in this chapter because the annihilation and generation of magnetic flux can be seen as two faces of related physical processes.

The primary goal of this book is to bridge the gap between the fundamental plasma physics and modern research on space storms. This is the challenge of the third part of the book. As in modern concertos, transition from the second to the third movement is not necessarily well-defined. In some sense Chapter 8 already opens the third part as here the treatise begins to focus more on the key issues in space storm research. Chapter 9, in turn, discusses the mechanisms giving rise to radiation that we see coming from the solar atmosphere at the time of solar storms as well as the scattering of radio waves from electrons and plasma fluctuations in the ionosphere. In Chapter 10 the adiabatic invariants introduced in Chapter 3 are used in formulating the kinetic equations for studies of plasma transport and acceleration in the inner magnetosphere.

Fluid turbulence remains one of the toughest problems in classical physics and turbulence in collisionless magnetized plasmas is an even harder problem. Particularly interesting environments, where turbulence is critical, are the interplanetary and planetary shocks with the associated sheath regions. Shocks and shock acceleration are discussed in Chapter 11.

Finally the treatise returns to the more phenomenological treatment of space storms in various parts of the solar–terrestrial system. Chapter 12 deals with the storms on the Sun and their propagation into the solar wind. In Chapter 13 magnetospheric storms and substorms and their drivers are investigated. As storm phenomena in the inner magnetosphere are of particular practical interest, they are discussed separately in Chapter 14. At the end of the journey some effects of space storms on the atmosphere and the current induction on the ground during rapid ionospheric disturbances are briefly discussed in Chapter 15.

The great variety of phenomena from the Sun to the Earth and the vast amount of different theoretical and modeling approaches to explain them make some hard choices necessary, in particular, the choice between a Sun–centered and an Earth–centered approach. The solar atmosphere, in particular the corona, is a much more stormy place than the Earth's environment. The Sun is also the driver of practically all space storm phenomena in the solar–terrestrial system. These facts would suggest adoption of the Sun–centered view on space storms. On the other hand, we live on the Earth and here we have to learn to handle the consequences of space storms. Thus the present choice is Earth-centered but more emphasis is put on the entire space storm sequence than in traditional textbooks on magnetospheric physics. There is a recent very comprehensive textbook on the physics of solar corona by Aschwanden [2004]. Actually just browsing through that volume, containing citations of about 2500 scientific articles, illustrates how difficult it is to compile a concise text on that end of the space storm chain. The first decade of the 21st century also forms a "golden age" of solar physics when several multi-wavelength spacecraft are producing an enormous amount of new empirical information on the active Sun. To digest all this will certainly take some time.

Another choice taken here is not to deal with space weather effects or practical modeling approaches. Concerning these we point the interested reader to the recent volumes by Bothmer and Daglis [2007] and Lilensten et al [2008] and references therein. In fact, the present book and those by Aschwanden [2004] and Bothmer and Daglis [2007] are strongly complementary to each other. They have quite different approaches but are dealing with closely related issues.

As one of the goals of this book is to provide material for advanced students, exercise problems of varying difficulty have been embedded within the text. They are grouped into three categories: Problems labeled *Train your brain* are mostly straightforward, often boring, derivations of expressions that are useful for students learning to master the basic material of the book. The label *Feed your brain* refers to problems or tasks that add to the reader's knowledge beyond the actual text and can also be useful for testing the reader's understanding of the material. Problems identified as *Challenge your brain* are a little harder (at least to the author), dealing also with unsolved or controversial issues. Creative solutions to some of these may be worth publishing in peer-reviewed journals.

A textbook discussing basic physics necessarily borrows material from earlier sources. The author was introduced to plasma physics through the classic texts by Boyd and Sanderson [1969], Krall and Trivelpiece [1973], and Schmidt [1979], which certainly can be recognized in the presentation of the fundamental plasma issues. When discussing "generally known" (or believed to be known) topics, in particular in Chapter 1, references to the scientific literature have been used sparsely. However, a number of references to some of the truly classic reports have been included. New generations of scientists every now and then tend to forget the original works with the risk of independent reinvention of the wheel. For students it is sometimes useful to recall that there was intelligent life even before they were born. In this respect the internet has actually made life much easier. We do no more need to have physical access to the best equipped libraries to read many of the classic reports in the scientific literature. Unfortunately, books like this are harder, or more expensive, to access electronically.

Acknowledgements

A large part of the material of this book comes from notes for space plasma physics, solar physics, and space weather lectures that I have been giving over the years to both master's and doctoral students, mostly at the University of Helsinki but also at several summer schools and other special occasions. I realized that there was a need for a book along the approach that I have taken, when I was leading a nation-wide space weather consortium in space research programme Antares of the Academy of Finland in 2001–2004. However, it was not until the academic year 2008–2009 that I was able to invest enough time in the project as the result of an appropriation for a senior scientist from the Academy of Finland, which facilitated a full year of sabbatical leave. I spent the autumn 2008 at the Laboratory for Atmospheric and Space Physics of the University of Colorado, Boulder, and the spring 2009 at the International Space Science Institute (ISSI) in Bern, Switzerland. I wish to express my sincere thanks to the directors, Dan Baker and Roger-Maurice Bonnet, and their staffs for the hospitality and support I received. Boulder provided an excellent academic environment for writing the main part of the text, whereas ISSI was the exactly right place for the hard work of editing and organizing the material.

Several people have influenced my thinking of space physics. Of my former mentors I wish to express my gratitude to Rolf Boström and Risto Pellinen, the latter of whom introduced me to the field of magnetospheric physics and supported me in many ways until his retirement. I am heavily indebted to the space physics community of the Finnish Meteorological Institute and the Department of Physics of the University of Helsinki. These two institutes and their close collaboration form a unique space research environment whose role in this exercise cannot be overestimated. In the context of the present book I wish most gratefully to thank Tuija Pulkkinen for excellent collaboration in research on storms and substorms over more than 20 years. Another person deserving special acknowledgment is Rami Vainio whose contributions to our space physics curriculum, in particular on the Sun and space plasma shocks, have been invaluable in writing the corresponding chapters of this book. Of my past and present local collaborators who have, explicity or implicitly, contributed to the book I wish to thank Olaf Amm, Natalia Ganushkina, Heli Hietala, Pekka Janhunen, Riku Järvinen, Esa Kallio, Kirsti Kauristie, Emilia Kilpua (née Huttunen), Tiera Laitinen, Jakke Mäkela, Anssi Mälkki, Heikki Nevanlinna, Minna Palm-

roth, Risto Pirjola, Antti Pulkkinen, Ilkka Sillanpää, Eija Tanskanen, Petri Toivanen, and Ari Viljanen. For technical help with scanning and editing of several figures I am grateful to Artturi Pulkkinen.

It is practically impossible to acknowledge all the colleagues whose ideas have somehow migrated into the the text. Both over the years and in the context of the present project I have had particularly useful collaboration and discussions with Mats André, Vassilis Angelopoulos, Dan Baker, Stas Barabash, Wolfgang Baumjohann, Joachim Birn, Joe Borovsky, Pontus Brandt, Jörg Büchner, Tom Chang, Eric Donovan, Lars Eliasson, Scot Elkington, Karl-Heinz Glassmeier, Georg Gustafsson, Gerhard Haerendel, Walter Heikkila, Bengt Holback, Gunnar Holmgren, Richard Horne, Mary Hudson, Bengt Hultqvist, Mike Kelley, Paul Kintner, Jim LaBelle, Xinlin Li, Mike Lockwood, Ramon Lopez, Bill Lotko, Tony Lui, Rickard Lundin, Larry Lyons, Bob Lysak, Göran Marklund, Bob McPherron, Tuomo Nygrén, Terry Onsager, Hermann Opgenoorth, Götz Paschmann, Tom Potemra, Geoff Reeves, Gordon Rostoker, Alain Roux, Ingrid Sandahl, Rainer Schwenn, Victor Sergeev, Jim Slavin, Yan Song, Rudi Treumann, and Don Williams.

Finally I wish to acknowledge the most helpful support provided by PRAXIS Publishing Ltd., in particular for reviewing by John Mason, copy-editing by Mike Shardlow, cover design by Jim Wilkie, and LaTeX help from Frank Herweg. I am extremely grateful to the publisher Clive Horwood for his enthusiasm and support during the process. I am particularly indebted to his patience when problems with my schedule after returning from the sabbatical led to a long delay with the delivery of the files for copy-editing.

Helsinki, September 2010 *Hannu E. J. Koskinen*

Units and Notation

SI units are used throughout the book. As a common exception energy and temperature are often expressed in electronvolts (eV), but in equations involving the temperature the Boltzmann constant k_B is written explicitly, in which case the temperature is given in kelvins (K). Furthermore, physical distance measures, such as the radius of the Sun ($R_\odot$), the radius of the Earth (R_E), or the astronomical unit (AU), are in frequent use. Also, when dealing with densities of a few particles per cm^3, or magnetic fields of a few nT, it is preferable to use these as units in order to avoid unnecessary use of powers of ten.

A person working within theoretical plasma physics or solar physics must also master the Gaussian cgs unit system, as much of the literature in these fields is still written in these units. Transformation from grams to kilograms, from centimeters to meters, or ergs to joules is trivial, but in formulas involving electrodynamic quantities the different unit systems are a nuisance. This sometimes leads to erroneous calculations, not only by factors of 10, but examples of errors by a factor of 3 or 4π are not too difficult to find in the literature, peer-reviewed articles included.

Macroscopic quantities in the three-dimensional configuration space are denoted by capital letters, e.g., electric current $\mathbf{J}$, fluid velocity $\mathbf{V}$, pressure P, etc., vectors in boldface and scalars in italics. The lowercase $\mathbf{v}$ is reserved to denote particle velocity as a function of time and the velocity coordinates in the phase space, e.g., in expressions as $f(\mathbf{r},\mathbf{v},t)$, whereas the lowercase $\mathbf{p}$ denotes the particle momentum $\mathbf{p}(t)$. In order to avoid conflict electric potential is denoted by φ, whereas ϕ is an angular variable. Similarly volume is denoted by $\mathscr{V}$ in order not to mix up it in some expressions with speed V. The volume differential in integral expressions is denoted by either d^3r or $d\mathscr{V}$.

In an ideal world a textbook should have a unique system of symbols. However, this is not a practical goal for a book that combines material from several different disciplines of physics, all with their own and by no means common or unique notations. Thus the most usual conventions are followed in the book, accepting that some symbols become heavily overloaded. One of them is μ, that in this book may denote the magnetic permeability of a medium, the magnetic moment of a charged particle, or the cosine of the pitch angle. J can denote the second adiabatic invariant, the absolute value of electric current $|\mathbf{J}|$, and omnidirectional particle flux. γ in turn appears as the polytropic index, as the Lorentz factor and in some instances as the wave growth rate, n as the particle density, the index of refraction and in vector form the unit normal vector, σ as electrical conductivity and the collision cross-section, etc. However, none of these ambiguities should lead to misunderstanding. After all, physicists are expected see the forest for the trees.

1. Stormy Tour from the Sun to the Earth

In addition to light and other wavelengths of electromagnetic radiation the Sun affects our environment through complicated plasma physical processes. The study of these interactions is known as solar–terrestrial physics. Already long before the space era there were indications that solar activity and geomagnetic perturbations must somehow be connected. A remarkable event was the large flare on the Sun observed, independently, by Carrington [1859] and Hodgson [1859] on September 1, 1859, after which a major magnetic storm commenced only 17 hours later. Today we understand that the storm was caused by a magnetic cloud associated with a coronal mass ejection (CME) that reached the Earth exceptionally quickly. The storm was very strong, evidently much stronger than any event recorded during the present era of space weather sensitive equipment in space and on the ground.

During the early 20th century the Sun was found to possess a highly variable magnetic field and the violent solar eruptions were found to somehow be related to strong magnetic variations observed on the Earth. But it was not until the dawn of spaceflight that the highly variable but continuously blowing solar wind was shown to be the agent that carries the perturbations from the Sun to the Earth. The variations in the solar wind shake the magnetic environment of the Earth, the magnetosphere. If the perturbations are strong enough, we call them "storms". We borrow terminology from atmospheric sciences and call the short-term variations in the solar–terrestrial system "space weather" and the longer-term behavior "space climate". In this book the term "space storm" is not limited to storms in the magnetosphere but includes stormy weather on the Sun, in the solar wind, and in the Earth's magnetosphere and ionosphere. Space storms at other planets form an interesting and intriguing complex of physics issues, the discussion of which, however, is beyond the scope of the present treatise.

1.1 Source of Space Storms: the Sun

Space weather and space climate are controlled by the temporal variability of the Sun in different time scales from minutes to millennia. In fact, when looking at the Sun with the

H.E.J. Koskinen, *Physics of Space Storms: From the Solar Surface to the Earth*,
Springer Praxis Books, DOI 10.1007/978-3-642-00319-6_1,

present observational tools, its surface and atmosphere are seen to be very stormy and noisy environments. In this section we review some of the basic properties of our active Sun. A modern introduction to the Sun itself is Stix [2002] and a wealth of material about the corona and its activity can be found in the comprehensive volume by Aschwanden [2004].

1.1.1 The Sun as a star

The physical picture of the Sun started to develop in the dawn of modern physical sciences when Galileo, one of the first developers and users of the telescope, observed sunspots on the solar disk. He showed in 1613 that they are structures on the surface of the Sun and not small planets as Schreiner had argued a few years earlier. After this promising start progress in solar physics remained slow. In 1802 Hyde discovered that solar spectrum contained several absorption lines, which were later cataloged by Fraunhofer. In 1844 Schwabe showed that the sunspot activity varies in an 11-year cycle and in 1859 Carrington and Hodgson observed a solar flare in white light. The second most common element in the universe was identified as late as 1868 in the solar spectrum by Lockyer and was later named helium.

Most of our present understanding of the Sun did not exist before the 20th century. Among the first major advances were Hale's measurements of intense magnetic fields in the sunspots in 1908, showing that whatever generated the solar activity, it was closely related to highly variable magnetism. An important enigma remained, however. In 1862 Sir William Thomson (later Lord Kelvin) had demonstrated that the largest imaginable energy source for solar radiation, the gravitational binding energy of the Sun, would not, at the present solar luminosity, be sufficient for more than 20 million years, which already at that time was considered far too short a history for the solar system. The solution to this problem required the development of quantum mechanics and finding of the nuclear forces. In 1938 Bethe and Critchfield described the dominant proton–proton reaction chain that powers the Sun. In this process 600 million tons of hydrogen is transformed to 596 million tons of helium, and the remaining 4 million tons is released as radiation.

After the revelation of nuclear fusion in the Sun an intensive puzzle work of fitting solar models to the increasing amount of detailed observation started with the goal of describing both the present structure and the past evolution of the Sun. From the mid-1970s the observations of solar oscillations and their interpretation, known as *helioseismology*, have become most important tools for reaching a very accurate description of the interior of the Sun.

Today we know that the Sun is a typical cool magnetic star. Its mass ($m_\odot$) is 1.99×10^{30} kg (330 000 times more massive than the Earth) and radius ($R_\odot$) 696 000 km (109 times the Earth's radius, R_E). The present Sun irradiates with a luminosity of 3.84×10^{26} W with an effective black body temperature of 5778 K. The Sun was formed about 4.55×10^9 years ago when an interstellar gas cloud with a mass of the order of 10^4 $m_\odot$ collapsed due to some interstellar gravitational perturbation, probably a shock wave, and further disintegrated, leading to the formation of the solar system. The collapse was not spherically symmetrical due to the presence of angular momentum and magnetic flux of the cloud.

While most of the angular momentum and magnetic flux were carried away by matter not ending up in the solar system, rotation and magnetic field are still today essential elements of the Sun and the solar system.

An intriguing obstacle on the road toward an acceptable solar model was the *solar neutrino problem*. Ever since the first neutrino experiments by Davis and Bahcall in the Homestake gold mine in 1967, observations based on different detection techniques indicated that the Sun would produce only 30–50% of the neutrino flux that the standard solar model predicts to arise from the fusion process in the core. Attempts to solve this problem, e.g., by adjusting the temperature of the central core, lowering the relative abundance of heavy elements, assuming a rapidly rotating core, or assuming a strong magnetic field in the core, all led to contradictions elsewhere in the solar models.

Meanwhile developments in neutrino physics started to point toward another solution based on the properties of the neutrinos themselves. Finally, strong evidence in favor of the nuclear physics explanation was obtained at the beginning of the 21st century with a Cherenkov experiment within a large water tank with a heavy water (D_2O) core at the Sudbury Neutrino Observatory [Ahmed and SNO Collaboration, 2004]. In that experiment it is possible to observe both the electron neutrinos, which are produced by the fusion, and the μ and τ neutrinos, to which a considerable fraction of the electron neutrinos are transformed through *neutrino oscillations* during the propagation from the Sun to the Earth

Figure 1.1 illustrates the main regions of the Sun (for a detailed discussion of the solar model, see Stix [2002]). The energy production takes place in the *core* within a radius of $0.25 R_\odot$ from the center of the Sun where temperature is 1.57×10^7 K and pressure 2.34×10^{16} Pa. From the core energy propagates outward through a very slow process of

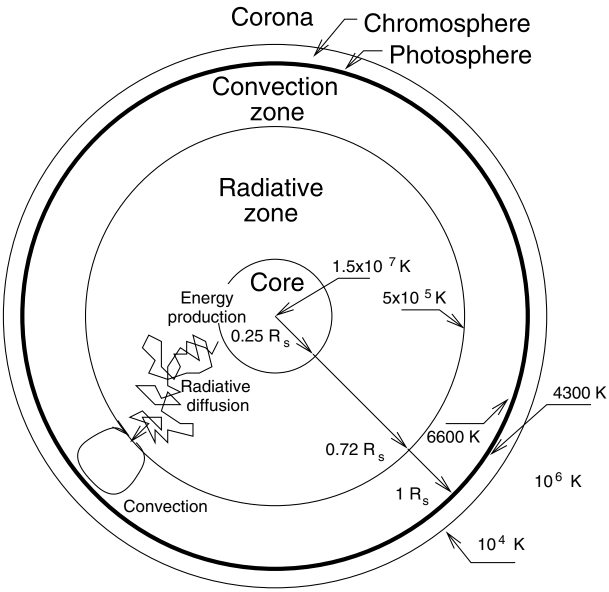

Fig. 1.1 The structure of the Sun. (Figure by courtesy of R. Vainio.)

radiative diffusion during which the photons are absorbed and re-emitted by the dense solar matter over and over again. The energy propagation time of the distance of 2 light seconds is of the order of 170 000 years. Due to collisions and absorption–emission processes in this *radiative zone* the photons are redshifted toward the visible wavelengths.

At the distance of about $0.72 R_\odot$ the solar gas becomes opaque to the photons and the energy transport toward the surface takes the form of turbulent convection, which is much faster than the radiative transfer. The plasma motion in this *convection zone* is extremely complex and of specific relevance to the topic of the present text, as the ever-changing magnetic field of the Sun is created within this zone, according to the present understanding close to its bottom. The radiation does not stop completely at the base of the convection zone. About $0.05\, R_\odot$ into the convection zone the convective energy flux exceeds the radiative flux and within the last $0.1\, R_\odot$ below the surface practically all energy transport is convective.

While the radiation zone is stably stratified, the convection zone is unstable: gas parcels move up, dissolve, and cool down, and the cool gas returns back along narrow lanes between the upward-moving gas parcels. The whole convection zone is continuously mixed, which makes it chemically homogeneous. This does not make the mean molecular mass constant because close to the surface the degree of ionization drops rapidly. However, within most of the convection zone the mean molecular mass is about 0.61.[1]

Finally the convection reaches the solar surface and introduces a granular structure on it. The intergranular lanes are about 100 K cooler than the regions of upward motion. Granules appear in various sizes, diameters ranging from about 1000 km up to a few times 10^4 km, the latter being called supergranules. The smallest granules represent small convection cells close to the surface, whereas the larger granules are related to larger convection cells reaching deeper into the convection zone.

Above the convection zone a thin surface, the *photosphere*, absorbs practically all energy carried by convection from below and irradiates it as (almost) a thermal black body at the temperature of 5778 K. The thickness of the photosphere is only 500 km. The temperature at the bottom of the photosphere is about 6600 K and at its top 4300 K.

The total irradiance at the mean distance of the Earth ($1\,AU$) is known as the *solar constant*

$$S = 1367 \pm 3\,\mathrm{W\,m^{-2}}\,. \tag{1.1}$$

It is related to the *luminosity* of the Sun $L_\odot$ by

$$L_\odot = 4\pi AU^2 S = (3.844 \pm 0.010) \times 10^{26}\,\mathrm{W}\,. \tag{1.2}$$

Accurate determination of S is challenging and the last digits and uncertainties in the expressions above must not be taken as definitive. The *total solar irradiance* (TSI) must be observed with accurately calibrated instruments above the dense atmosphere, which absorbs most of the radiation in ultraviolet (UV) and infrared (IR) wavelengths. Early in the 21st century a consensus of inter-calibrations between various space observations was reached of an average $S \approx 1366\,\mathrm{W\,m^{-2}}$ near solar minima and $S \approx 1367\,\mathrm{W\,m^{-2}}$ near

---

[1] In a plasma free electrons are counted as particles. Thus the mean molecular mass of electron–proton plasma is 0.5.

solar maxima. However, observations with the Total Irradiance Monitor (TIM) onboard the *The Solar Radiation and Climate Experiment (SORCE)* satellite launched in 2003 indicate that the actual TSI would be some 4–5 $\mathrm{W\,m^{-2}}$ smaller than previously thought [Kopp et al, 2005]. By the time of writing this book the reason for this discrepancy had not been clarified.

For space storms the exact total irradiance is not as important as its relative variations. In particular, near solar maxima the irradiance varies by several $\mathrm{W\,m^{-2}}$ depending on the sunspot activity (Sect. 1.1.5).

The luminosity can be given in terms of the *effective temperature* defined by

$$L_\odot = 4\pi R_\odot^2\, \sigma T_{\mathit{eff}}^4 \,, \tag{1.3}$$

where $\sigma = 5.6704 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$ is the *Stefan–Boltzmann constant*. The effective temperature of the Sun is $T_{\mathit{eff}} = 5778 \pm 3\,\mathrm{K}$. The photospheric gas has this temperature at the optical depth $\tau \approx 2/3$, which can be taken as the definition of the solar surface (for the definition of τ, see, e.g., Stix [2002]).

"Solar constant" is actually one of many historical misnomers that we will encounter in this book. The Sun is a variable star in both short and long time scales. Fortunately for us, the variations are about a factor of three weaker than is typical for many other Sun-like stars. In the longest time perspective the luminosity of the newly-born Sun was about 72% of its present value. After some 2 billion years from now the Sun will have become so bright that the Earth will turn too dry for the present type of life. The slow rise of solar luminosity is due to the increase of the core temperature when more and more hydrogen is fused to helium.

In space weather and space climate time scales, S varies by a factor of

- 10^{-6} over minutes
- 2×10^{-3} (0.2%) over several days
- 10^{-3} over a solar cycle (the number is quite uncertain because the solar cycles are different)

The physical reasons and apparent periodicities for these variations are not fully understood.

1.1.2 Solar spectrum

The solar spectrum from γ-rays to metric radio waves is given in Fig. 1.2. Most of the solar energy is irradiated in the visible and near-infrared parts of the spectrum with peak irradiance in yellow light around 450–500 nm. The red end of the spectrum is an almost continuous black-body spectrum with some strong absorption lines, e.g., Hα at 656.3 nm (not visible in the scale of Fig. 1.2). At the blue end there are more absorption lines.

About 44% of the electromagnetic energy is emitted at infrared wavelengths $\lambda > 0.8\,\mu\mathrm{m}$. This part of the spectrum is approximately thermal and can be represented by the *Rayleigh–Jeans law*

$$S(\lambda) \simeq 2ck_BT\,\lambda^{-4}\,(R_\odot/AU)^2 \,. \tag{1.4}$$

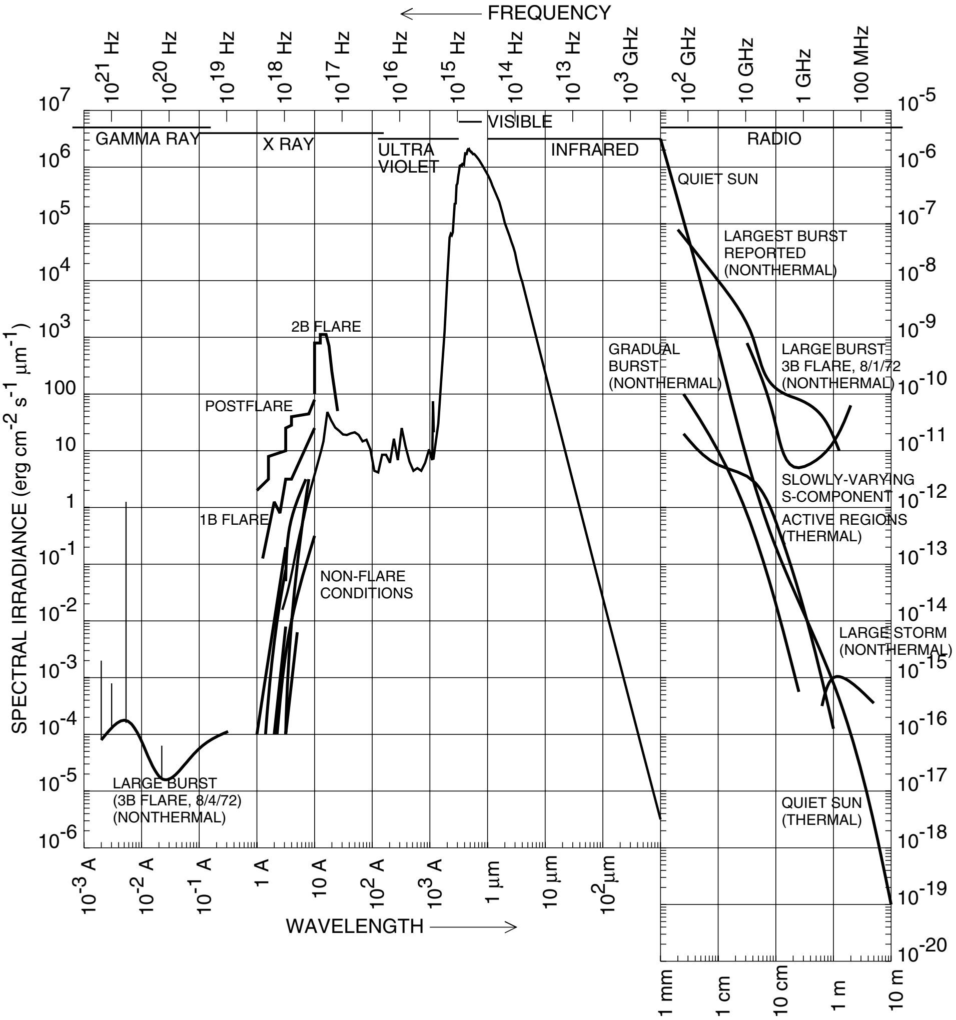

Fig. 1.2 Solar spectrum from γ-rays to radio waves. The radio wave part of the spectrum is shifted up in irradiance by 12 orders of magnitude. The irradiance is given in cgs units and ångström (1 Å = 0.1 nm) is used below one 1 μm, which is common practice in solar physics. (From Aschwanden [2004].)

The infrared spectrum is absorbed mostly by water vapor in the Earth's atmosphere.

At radio wavelengths (> 1 mm) the spectrum is commonly presented as a function of frequency (recall the conversion: $\lambda(\mathrm{m}) = 300/f(\mathrm{MHz})$; e.g., 1 mm $\leftrightarrow$ 300 GHz). The Sun is strongly variable at these wavelengths because the radio emissions originate from nonthermal plasma processes in the chromosphere and corona (discussed in Sect. 1.1.3). As indicated in Fig. 1.2, the radio emissions during strong solar storms can exceed the quiet levels by several orders of magnitude. Note that there is an ankle in the slope of the quiet-Sun spectrum at around 10 cm indicating higher temperatures ($\sim 10^{6}$ K) than the main black body radiation. This is a signature of the chromosphere and corona being much hotter than the visible Sun.

In the ultraviolet side of the spectrum absorption lines are dominant down to 210 nm. At shorter wavelengths the intensity is reduced to correspond to the temperature of 4700 K. This reduction is due to absorption by the ionization of Al I. (Recall the notation: Al I represents non-ionized aluminum, Al II is the same as Al^+, Al III is Al^{2+}, etc.) Below 150 nm emission lines start to dominate the spectrum. The strongest is the hydrogen Lyman α line centered at 121.57 nm. Its average irradiance, 6 mW m^{-2}, is as strong as all other emissions below 150 nm together and the line is also clearly visible in Fig. 1.2 .

At shorter wavelengths the spectrum becomes highly variable, illustrating a nonuniform distribution of the emission sources in the solar atmosphere. The nonuniformity is both spatial and temporal. The wavelength band below 120 nm is called *extreme ultraviolet* (EUV). These emissions come both from neutral atoms and from ions up to very high ionization levels, e.g. Fe XVI (Fe^{15+}) in the solar corona. This facilitates the observations of the wide range of temperatures from 8000 K to 4×10^6 K, from the chromosphere to the corona.

Solar flares increase the EUV and soft X-ray (0.1–10 nm) spectra quite considerably. Also hard X-rays and γ-rays are emitted in these processes, as will be discussed in Chap. 12.

1.1.3 Solar atmosphere

That there is an atmosphere above the photosphere is evident already visually. The irradiance decreases from the center of the disk to the limb by an order of magnitude due to the absorption of the atmospheric gas, which is known as *limb darkening*. The temperature continues to decrease in the photosphere reaching its minimum at an altitude of about 500 km. Thereafter, the temperature starts to rise again in the *chromosphere*. The chromosphere has got its name from the colorful flash seen just at the beginning and at the end of a total solar eclipse. The most prominent color is the red Hα-line at 656.3 nm. Traditionally the chromosphere was thought to be a layer of thickness of about 2000 km, but as illustrated in Fig. 1.3 the present view to the structure of the solar atmosphere is much more complicated and dynamic than the old picture of a gravitationally stratified atmosphere.

At the upper end of the chromosphere the temperature begins to rise more rapidly. The chromosphere is sometimes defined to end at the temperature of 25 000 K. Above the chromosphere there is a thin *transition region* to coronal temperatures of the order of 10^6 K. The *corona* is a key region of many aspects of space storms to which we will return in Sect. 1.1.6.

The steep temperature increase from the chromosphere to the corona remains one of the major insufficiently understood topics in solar physics. As illustrated in Fig. 1.3 the chromospheric and coronal plasmas partly overlap, flowing up and down with complicated dynamic magnetic field structures involving waves, shocks, magnetic reconnection, etc., which will be discussed in later chapters of this book. At the same time when this dynamism complicates the picture, it also indicates that there free energy is available for the heating. In fact, a steep temperature gradient in a gravitationally stratified atmosphere

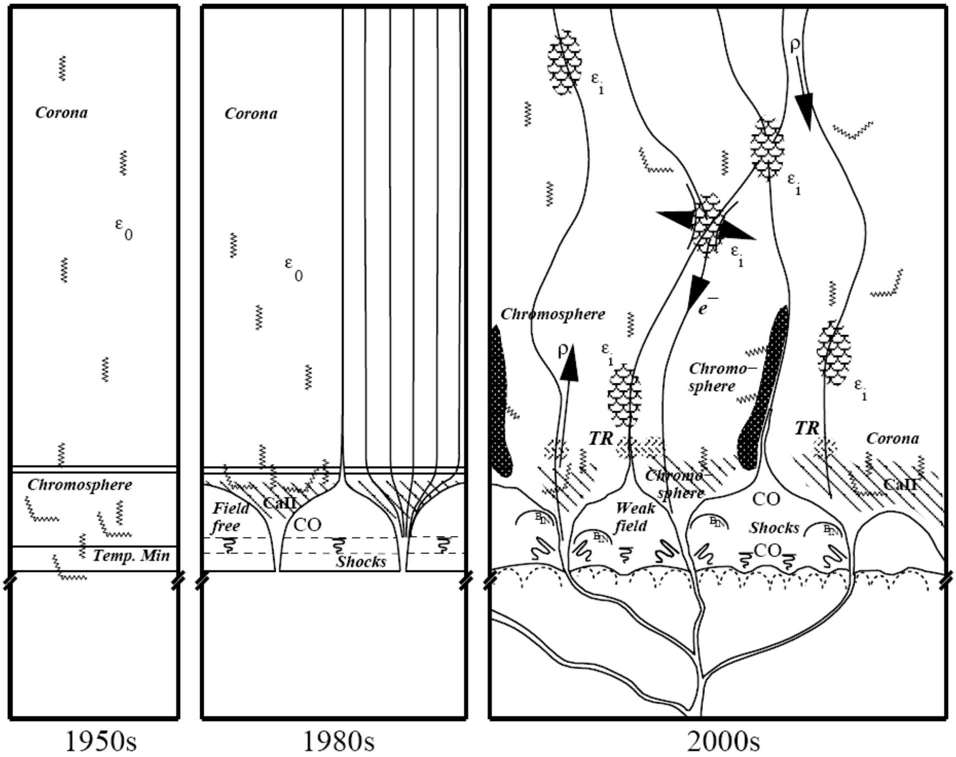

Fig. 1.3 Evolution of the concepts the solar atmosphere from gravitationally stratified layers in the 1950s to a highly inhomogeneous mixing of the photosphere, chromosphere, and corona at the beginning of the 21st century. (From Schrijver [2001].)

might be much more difficult to explain than a spatially and temporally variable environment.

1.1.4 Rotation of the Sun

That the Sun rotates was discovered soon after the advent of telescope in about 1610. Around 1630 it became clear that the rotation is not rigid, but the equatorial surface rotates faster than the high-latitude regions. The origin of this *differential rotation* is not yet fully understood. It is related to the transport of angular momentum inside the Sun and it also plays a central role in the generation of the solar magnetic field. Differential rotation appears to be a general property of self-gravitating large gaseous bodies and is also observed in the giant planets of the solar system.

The rotation axis of the Sun is given by two angles: the *inclination i* between the ecliptic plane and the equatorial plane, and the *angle of the ascending node α* of the Sun's equator, i.e., the angle in the ecliptic plane between the direction of the vernal equinox and the direction where the solar equator cuts the ecliptic from below. The Earth's precession shifts the equinox direction by 0.0196°, i.e., $50''$, per year, and thus α increases by the same rate. Consequently, the *epoch* must be given when coordinates related to the equinox are used. Carrington determined these angles in 1863 as $i = 7.25^\circ$ and $\alpha(1850) = 73.67^\circ$. The latter is still valid but the Greenwich sunspot data from the period 1874–1976 imply $i = 7.12^\circ \pm 0.05^\circ$.

We denote the *heliographic latitude* by ψ, thus the polar angle (co-latitude) is $\theta = \pi/2 - \psi$. There is no physically unique way to define the longitude on the differentially rotating surface. For this purpose Carrington introduced a notation that is still in use. He divided time into intervals of 27.2753 days. These intervals are called *Carrington rotations*. Carrington rotation 1 was defined to have commenced on 9 November 1853. In one year of 365 days there are 13.38 Carrington rotations and thus the present rotation numbers are well over 2000. At the commencement of a new rotation longitude $\phi = 0$ is attached to the center of the solar disk. Note that the Carrington rotations are related to the motion of the Earth around the Sun, i.e., the "same place" at the solar equator is toward the Earth after one Carrington rotation. This is known as the *synodic period*. The "true" rotation period with respect to the stars is the *sidereal period* of about 25 days.

Carrington determined the surface rotation rate from sunspot data as a function of the heliographic latitude in (sidereal) degrees per day

$$\Omega(\psi) = 14.25 - 2.75 \sin^{7/4} \psi \, . \tag{1.5}$$

The power 7/4 is a bit awkward. A more modern approach is to expand the rotation rate as

$$\Omega(\psi) = A + B \sin^2 \psi + C \sin^4 \psi + \ldots \tag{1.6}$$

and in most studies only coefficients A and B are determined. Here A is the equatorial rotation rate.

In addition to sunspot data, Doppler shifts, edges of coronal holes and surface magnetograms are used in studies of the rotation rate. The different methods yield slightly different results and there is some variability within the individual methods as well. Furthermore, different sunspot cycles are different. For example, Pulkkinen and Tuominen [1998] used the sunspot data from cycles 10–22 (years 1853–1996) and found that the coefficients varied in the ranges $A = (14.38\,, 14.85)$ and $B = (-3.19\,, -2.51)$.

It is interesting to note that the larger the structure used to determine the rotation, the more uniform rotation is found. The extreme are observations of large coronal holes, which sometimes show very little differential rotation at all. During the last decades helioseismology has revolutionized the studies of differential rotation. Now it is possible to empirically determine the rotation also inside the Sun, as illustrated in Fig. 1.4, which has been derived from the observations of solar oscillations using the MDI instrument onboard the *SOHO* spacecraft.

A rotating non-rigid body is not fully spherical. Even the Earth is elastic and has an *oblateness* $f = (r_{eq} - r_{pol})/r_{eq} \approx 1/300$. The fast-rotating gas giant planets Jupiter and Saturn are much more oblate, $f_J = 0.065$ and $f_S = 0.098$, which can be perceived already in rather low-resolution pictures. But how oblate is the slowly rotating Sun, whose exact diameter is difficult to measure?

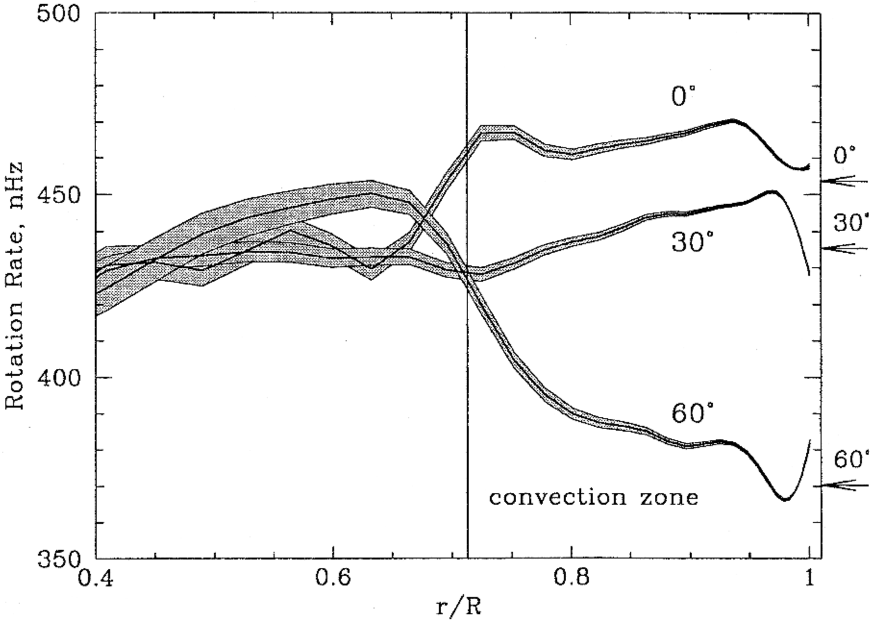

Fig. 1.4 The internal rotation rate of the Sun. The radial profiles are calculated for three different latitudes. The grey regions indicate the estimated error in the inversion procedure. (From Kosovichev et al [1997].)

Neglecting the differential rotation and expanding the external gravitational field up to the quadrupole term (the first non-zero correction)

$$\Phi_{ext} = -\frac{Gm_{\odot}}{r}\left[1 - J_2\left(\frac{R_{\odot}}{r}\right)^2 P_2(\theta)\right] \tag{1.7}$$

the oblateness expressed as $\triangle r/R_{\odot}$ is

$$\frac{\triangle r}{R_{\odot}} = \frac{1}{2}\frac{\Omega^2 R_{\odot}}{g_{\odot}} + \frac{3}{2}J_2 , \tag{1.8}$$

where Ω is the angular velocity of the solar surface, J_2 the quadrupole moment and $P_2(\theta)$ the second Legendre polynomial. Using the Carrington rotation rate, the first term in (1.8) is about 10^{-5}.

In the past the Sun has rotated faster than today. The specific angular momentum (i.e., the angular momentum per unit mass) of the cloud collapsing to form the Sun was much larger than the angular momentum of the present solar system. Much of this was lost in a very early phase of the solar evolution. We know that the so-called *T Tauri stars*, which are in the early phase of their evolution, rotate much faster than the Sun. Their surface velocities are about 15 km s^{-1} compared to 2 km s^{-1} of the present Sun.

According to pre-main-sequence stellar models, the Sun was fully convective before the hydrogen burning started. The convection was turbulent and the rapid exchange of momentum between parcels of gas evened out the gradients in the angular velocity. The total angular momentum J_0 has been estimated to have been $8 \times 10^{42}\,\mathrm{kg\,m^2\,s^{-1}}$, whereas it presently is $1.7 \times 10^{41}\,\mathrm{kg\,m^2\,s^{-1}}$.

Matter leaving the Sun carries angular momentum, but the material loss since the time of large J_0 has been negligible. The magnetic field, however, is a very efficient lever arm for a torque. As we will discuss in the context of the solar wind (Sect. 1.2.2), the magnetic field forces the escaping material to rotate with the Sun out to the so-called *Alfvén radius* $r_A \approx 12 R_\odot$. Thus the angular momentum density increases up to r_A, and it is this angular momentum that is conserved in the escaping flow beyond r_A. The rate of angular momentum loss is

$$\frac{dJ}{dt} = \Omega\, r_A^2 \,\frac{dm}{dt}\,. \tag{1.9}$$

How much such *magnetic braking* really has taken place in the history is difficult to estimate because we do not know the history of the magnetic field on which r_A depends. The magnetic field is generated by the solar dynamo (Sect. 8.3.2), which depends on Ω and in particular on its gradient. As long as the Sun was fully convective the slowing down affected the whole Sun. When the radiative core developed, the motion of the outer convective zone was disconnected from the interior. The convective part continued to lose angular momentum by magnetic braking, but what happened to the core? Because the central core contracted further, the first guess would be that its rotation rate should have increased.

However, the recent results of helioseismology (e.g., Fig. 1.4) do not support the idea of a fast-rotating core. The central core may rotate somewhat faster than the radiative zone but something seems to have slowed down the rotation also in the inner parts of the Sun. A strong inward gradient $d\Omega/dr$ would mean strong shear flows. These could drive instabilities, which, in turn, could transport the excess angular momentum, resulting in smoother $d\Omega/dr$. It has also been speculated that there could be an internal magnetic field in the core. Indeed, already a relatively weak magnetic field would be sufficient to slow down the core.

1.1.5 Sunspots and solar magnetism

The magnetic field of the Sun is very complicated both in time and in space. The existence of solar magnetic fields was first found in *sunspots* by Hale in 1908. Although we can today measure much weaker magnetic fields on the Sun, the sunspots have retained a central role in studies of solar magnetism. The theory of magnetic field generation is a difficult topic of plasma physics, and after a century of intensive study we still lack a fully satisfactory physical description of the generation and evolution of the solar magnetic field.

A sunspot corresponds to an intense magnetic flux tube emerging from the convection zone to the photosphere. Large spots can have diameters of about 20 000 km. The center of the spot is called the *umbra* whose temperature is about 4100 K, and the largest observed magnetic fields are about 0.3 T. The strong magnetic field is the cause of the low