SPRINGER BRIEFS IN EARTH SCIENCES

Erdal Yiğit

Atmospheric and Space Sciences: Ionospheres and Plasma Environments

Volume 2

Springer

SpringerBriefs in Earth Sciences

More information about this series at http://www.springer.com/series/8897

Erdal Yiğit

Atmospheric and Space Sciences: Ionospheres and Plasma Environments

Volume 2

Springer

Erdal Yiğit
Department of Physics and Astronomy
George Mason University
Fairfax, VA
USA

ISSN 2191-5369 ISSN 2191-5377 (electronic)
SpringerBriefs in Earth Sciences
ISBN 978-3-319-62005-3 ISBN 978-3-319-62006-0 (eBook)
DOI 10.1007/978-3-319-62006-0

Library of Congress Control Number: 2017944208

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my beloved parents
İhsan and Nazlı Yiğit and to all children of the universe.

Foreword

For me as an early participant in space research it was mandatory to study the ionosphere, since it was in this environment that our experiments with plasma clouds were performed. The textbooks available at that time emphasized production and loss of ions and electrons by photoionization and particle collisions, i.e. local processes. Rocket experiments with neutral gas releases and plasma cloud injections showed, however, that coupling between regions was essential for understanding our results. The same applied to the rising research with incoherent scatter radar. On the one hand, there was besides particle precipitation the downflow of momentum and electromagnetic energy and the closure of electric currents from the magnetosphere. On the other hand, a feedback from the ionosphere was recognized as controlling magnetospheric dynamics and structuring. While ionospheric physics was greatly enriched by the ongoing research, it was found convenient that the ionosphere could at least be treated as if having an impenetrable lower boundary at the base of the D-region. It was only through my colleagues working with ionospheric radars that I learnt about gravity waves which might propagate into the ionosphere and cause sporadic-E. One of the great virtues of this monograph is that it emphasizes these so far missing coupling aspects from below, mainly by waves and circulation, a field to which the author has made important contributions. Far from covering the broad range of plasma phenomena it offers, on the one hand, a fresh approach to the fundamentals of the plasma dynamics and, on the other hand, enriches the descriptive sections by a great many citations of contemporary research. Thereby it offers to the newcomer of the field acquisition of solid foundations and plenty of stimulation for further in-depth studies. A third and very timely virtue of this monograph is that the author always keeps the eyes of the reader on the application of the physical relations to the modelling of planetary atmospheres and ionospheres and space weather effects. This is of particular value at a time where modelling has increasingly acquired the aspect of creating a world parallel to lab and space with experiment and diagnostics.

Having met the author sixteen years ago as one of the pioneering students at the International University Bremen, now Jacobs University, I want to congratulate him for having completed already two monographs, whereas I in my long life have not

even managed to deliver one. So I conclude with the hope that he may one day undertake writing a third monograph that is badly missing, namely on the coupling of solar wind and magnetosphere with the ionosphere and atmosphere. Gerhard Haerendel

Garching
April 2017

Prof. Dr. Gerhard Haerendel

Preface

It has been a great pleasure to successfully complete the first volume of my monograph, the "Atmospheric and Space Sciences: Neutral Atmospheres", in which I have focused on basic atmospheric physics and neutral atmospheric processes. Earth's atmosphere is essentially a thin layer of gas around a solid planet, much like the thin skin around an onion. My goal was and still is to help bridge the gap between historically separated fields of meteorology, space science (sometimes called aeronomy), and planetary sciences. At least I would like to demonstrate how these fields are interconnected and similar type of basics physics is applicable to address the different processes these fields study. The atmosphere consists of several gas layers that are distinguished from each other by the differences in the properties of physical and chemical mechanisms. When I was a Ph.D. student, I remember having reflected on this division and realized how challenging it can be for the members of these fields to communicate with each other about the extent to which different processes are important in the atmosphere. To illustrate this problem of communication I can remember the "Chapman Conference on Atmospheric Gravity Waves and their Effects on General Circulation and Climate" that took place in Honolulu, Hawaii in February 2011. I gave a talk on gravity waves (GWs) with the title "Dynamical and Thermal Effects of Gravity Waves in the Terrestrial Thermosphere-Ionosphere". It was the only conference talk on the direct effects of lower atmospheric GWs on the upper atmosphere. After my talk during the discussion period, one of the scientists had asked me why one would study wave effects in the upper atmosphere. His rationale for this question was: "There is anyway nothing up there in the thermosphere"…

So, the second volume, the "Atmospheric and Space Sciences: Ionospheres and Plasma Environments" will focus primarily on planetary ionospheres and the basic electromagnetic processes (i.e., plasma physics) that govern them. I have retained the focus on Earth's ionosphere, as I had focused more on Earth's neutral atmosphere in the first volume. Often I will use the expressions of "ionospheres" and "planetary ionospheres" interchangeably. I would like to continue the tradition that I have established in the first volume of including basic physics as well as research topics in the same volume. By no means the selected topics cover all aspects of

atmospheric and space sciences. There are many other interesting topics out there that deserve to be included in a more comprehensive text book or monograph and are out of the scope of my book.

Meanwhile, after the publication of the first volume, I have talked to several research colleagues, graduate and undergraduate students on the philosophy of this monograph series. I had the impression that I had, at least, partially achieved my goal of providing a different perspective on atmospheric and space sciences. In general, a rich collection of references have been used in the book, which should serve the reader as an opportunity to investigate some of the details of various research topics. It is my humble intention that the reader will be inspired by the collection of the discussions and different topics in the text or by the art of presentation.

The selection of research topics in the book is not necessarily thought to indicate an evaluation of the significance of these topics with respect to many other topics that I have not covered. My selection only reflects a subjective view of a collection of important research topics. Nevertheless, I find that meteorological effects on the upper atmosphere, which we term as "coupling from below", and "space weather" effects, referred to as "coupling from above", are two highly important rapidly rising areas of research in the atmospheric and space sciences community. Having personal interest and scientific contributions in these topics, I felt obliged to provide some special coverage in this book. I think, various research activities in these two fields make up a significant portion of the recent activities and progress in atmospheric and space sciences. Nevertheless, atmospheric coupling and space weather research provide an excellent motivation toward unifying atmospheric and space physics.

Fairfax, VA, USA Erdal Yiğit
May 2017

The original version of the book was revised: author name has been corrected. An erratum to this book is available at https://doi.org/10.1007/978-3-319-62006-0_6

Acknowledgements

The science world has provided a large pool of papers that have helped me complete the second volume of my monograph. Thus, I cannot thank enough to all authors of basic ionospheric physics papers published earlier. Essentially, ionospheric physics is about a century old and there have been many interesting contributions that I have not had the space to acknowledge here.

I am very much indebted to Petra van Steenbergen, Springer Senior Publishing Editor for Earth Sciences and Geography, for not losing her faith in me in terms of the completion of the book. I have passed the submission deadline several times and she was very open-minded to all the challenges I have had. It was also a great pleasure for me to chat with her about general science.

Dr. Medvedev has continuously interacted with me and inspired some parts of the book. Overall, he was very tolerant of me when I did not have enough time to work on the urgent projects. I always appreciated that he has intrinsically assumed I must be busy with the book and have avoided to "push me" hard.

I would like to express my special thanks to Prof. Dr. G. Haerendel, who had kindly agreed to write a foreword for this monograph. As an enthusiastic scientist, he has been an excellent role model during my physics degree.

The greatest energy came from my parents whose existence I am very much grateful for. Although my father does not have expertise in space science, he has always asked curious and technical questions.

Contents

Acronyms

CHAMP	CHAllenging Minisatellite Payload
CME	Coronal Mass Ejection
DMSP	Defense Meteorological Satellite Program
EM	Electromagnetic
ESA	European Space Agency
EUV	Extra Ultraviolet
GCM	General Circulation Model
GITM	Global Ionosphere Thermosphere Model
GNNS	Global Satellite Navigation System
GRACE	Gravity Recovery and Climate Experiment
GW	Gravity Wave
IAGA	International Association of Geophysics and Aeronomy
ICMA	International Commission in the Middle Atmosphere
IMAGE	Imager for Magnetopause-to-Aurora Global Exploration
IMF	Interplanetary Magnetic Field
IRI	International Reference Ionosphere
IUGG	International Union of Geodesy and Geophysics
MAVEN	Mars Atmosphere and Volatile EvolutioN
MHD	Magnetohydrodynamic(s)
NASA	National Aeronautics and Space Administration
NGIMS	Neutral Gas and Ion Mass Spectrometer
QBO	Quasi-biennial Oscillation
SCOSTEP	Scientific Community on Solar-Terrestrial Physics
SORCE	Solar Radiation and Climate Experiment
SSI	Solar Spectral Irradiance
SSW	Sudden Stratospheric Warming
TGO	Trace Gas Orbiter

TIMED	Thermosphere Ionosphere Mesosphere Energetics and Dynamics
TIM	Total Irradiance Monitor
TSI	Total Solar Irradiance
UT	Universal Time
UV	Ultraviolet

Chapter 1
Introduction to Plasma

History and Fundamental Parameters

The Earth is the cradle of humanity, but mankind cannot stay in the cradle forever. Man must at all costs overcome the Earth's gravity and have, in reserve, the space at least of the Solar System.

Konstantin Tsiolkovsky (1857–1935)

Abstract A plasma is an electrically conducting quasi-neutral gas mainly composed of charged particles that exhibit collective motion. The vast majority of observable universe consists of plasma. Closer to Earth, i.e., in the solar-terrestrial environment, the solar wind, the magnetosphere, and the ionosphere are the most studied plasma environments. Plasma processes are greatly influenced by ambient electromagnetic fields and, in analogy with neutral planetary atmospheres, waves are continuously generated in plasmas and their propagation and interaction influence the structure and the evolution of plasma. A naturally occurring magnetic field environment is the geomagnetic field, which has already been detected four centuries ago. In this introductory chapter, a conceptual discussion of plasma is first presented. Then, various space plasma environments are discussed along with a brief introduction to the history of space research. Finally, some key plasma parameters such as density, gyrofrequency, and temperature, and characteristics (collective motion, quasi-neutrality) are discussed concisely.

Keywords Plasma · Plasma parameters · Space plasma environment · Ionosphere · Magnetism · Electromagnetic field · Collective motion · Quasi-neutrality · Planetary missions

1.1 What Is Plasma?

The word "plasma" originates from the Greek word πλασμα, which means a shape or structure. Plasma is an ionized substance, that is, it contains charged particles (electrons, ions) that can move freely. It is in fact the most abundant state of matter in the universe, although we are not very often directly concerned with it in our

E. Yiğit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*, SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-62006-0_1

day-to-day life. Sometimes, it is referred to as the fourth state of matter, besides the liquid, gas, and solid states. The rationale for this is that plasma demonstrates complex behavior that cannot be simply described considering only one of the three states. Typically the charged particles that constitute the plasma are at high temperatures and have relatively large densities. Plasma can be produced in a laboratory environment by heating a gas to a very high temperature or by generating an electric discharge within the gas. The subject of plasma physics is dedicated to the study of the behavior and properties of plasma. The outer space contains various forms of plasma. Though, we do not need to travel that far away to encounter plasma. Fluorescent lamps, which are more efficient than incandescent lamps, contain electrically conducting plasma gas. One of the most obvious atmospheric plasma phenomena is lightning, which is a naturally occurring atmospheric electric discharge. It occurs because of charged particle (plasma) flow between a cloud and surface. Consider a discharge lamp, in which positively charged ions are generated during high-energy collisions and form plasma together with the negatively charged electrons. So, when you are looking at a red shining neon sign, you are actually looking at plasma. Overall, from Earth's inner core to the upper regions of the atmosphere, plasma can be found all around. Above Earth's surface, starting around 70–80 km, we have a consistent layer of plasma, which forms the ionosphere. Other known planetary atmospheres have well-developed ionospheres as well. Our Sun, which enables the existence of life, is a giant nuclear fusion reactor in which high-density plasma is contained and the associated fusion energy is transported from its core to the surface, where it is radiated as electromagnetic energy. Even in the interstellar medium, i.e., in the regions between galaxies, there is plasma.

One of the foci of this book is to provide an introduction to ionospheric physics, which is in essence the study of the physics and chemistry of atmospheric plasma and its interaction with the neutral atmosphere. Overall, we have more than enough reasons to be curious about plasma and a formal definition of plasma is well-deserved in this context. A plasma is an electrically conducting *quasi-neutral* gas mainly composed of charged particles that exhibit *collective motion*.

The communication technology is exploting the plasma properties of the ionosphere, one of the most prominent plasma environment on Earth. In order to better appreciate what a plasma is let us first discuss two fundamental plasma properties mentioned above, the quasi-neutrality and collective motion, in further detail. Some behavior of matter in nature can be understood in term of its electric charge, which is controlled by the distribution of charged particles and is generally characterized by the number density of various charged particles, i.e., number of charged particles per unit volume [m^{-3}]. Charged particles are electrons, which have small mass and are negatively charged, and ions, which are much heavier than electrons and can be negatively or positively charged. Therefore ions occur because of charge imbalance. If an atomic or molecular species contains more electrons than protons, then it is a negatively charged ion. If there are more protons present than electrons then it is a positively charged ion. Ions are much heavier than electrons because they contain at least one or more protons, where a proton mass is about 1836 times larger than an electron mass, i.e., $m_p \gg m_e$ thus in most calculations of electrodynamics,

terms including the electron mass can be neglected safely with respect to the ones including the proton/ion mass. Note that the mass of an electron is tiny:

$$m_e = 9.109 \times 10^{-31}\ \text{kg}. \tag{1.1}$$

The masses of a proton and a neutron are very similar, but the neutron is about 0.12% larger than the proton: $m_n = 1.675 \times 10^{-27}$ kg versus $m_p = 1.673 \times 10^{-27}$ kg. In comparison the weight of a single dust particle is in the order of $\sim 10^{-13}$ kg. Such a small weight is hard to imagine.

Consider the simplest hydrogen ion $^1H^+$, one of the most abundant ions in the universe, as an example. It is essentially a proton as it consists of one proton only, thus $m(^1H^+) \gg m_e$. Hydrogen plays an important role in the Solar System. Fusion of four hydrogen nuclei to Helium ($^4_2$He), during which large amounts of energy is released, fuels Sun's energy, which is essential to the common forms of life that exist on Earth. Two hydrogen atoms are basic building blocks of a water molecule, H_2O. Neutrals are by definition not charged. For example, the oxygen atom ($^{16}_8$O), which is the most abundant species in the terestrial upper thermosphere, is a neutral particle.

An elementary charge e has a value

$$e = 1.602 \times 10^{-19}\ \text{C} \tag{1.2}$$

and has a negative value for an electron, while ions can be positively or negatively charged and one can thus have $\pm e$. It is a spectacular property of the microscopic world, i.e., the tiny world of atoms and molecules that we cannot observe directly with the human eye, that electrons and protons have the same magnitude of charge but greatly different masses. I find this property quite fascinating but this aspect is better discussed probably in the context of particle physics or even philosophy.

Coming back to the technical aspects, let us further discuss what controls plasma. Electric fields (**E**) and magnetic fields (**B**) together, i.e., electromagnetic fields (EM-fields), greatly influence the properties and motion of the plasma. When charged particles move randomly in space, local effects occur, leading to local formation of negatively and positively charged regions, which give rise to electric fields. Moving charges give rise to currents and magnetic fields. Generation of currents can take place in microscopic or macroscopic scales. For example, the motion of Jupiter's moon Io, the volcanic "moon", around Jupiter sets up complex current systems within Io, which initiates powerful volcanic activity on Io's surface, which is a fascinating manifestation of large-scale current systems. In the ionosphere, small-scale wave motion generate currents with high temporal variability. Also, Earth's ionosphere and magnetosphere systems are tightly coupled to each other by large-scale current systems. Specifically, magnetospheric physics deals with the mathematical theory of plasma dynamics considering self-consistent coupling between the magnetosphere and ionosphere. This coupling produces complex patterns of magnetospheric convection at auroral and low-latitudes, which responds to variations in changing orientation of the interplanetary magnetic field (Vasyliunas 2010).

Overall, fields and charged particles are tightly coupled to each other. However, the interaction between plasma flows and (electric) fields have sometimes been heavily discussed due to a lack of consensus on the understanding of their interactions (e.g., Vasyliunas 2001). The generated fields influence the motion of other charged particles remotely. This indirect coupling within the plasma is understood as a collective behavior. Thus, for the description of plasma dynamics, electromagnetic forces have to be accurately represented. The quasi-neutrality means that the number of negatively charged particles (electrons) are equal to the number of positively charged ions, $N_e = N_i$, in a given volume and thus the number density of electrons is equal to the number density of ions, i.e., $n_e = n_i$. In a macroscopic scale, the plasma behaves as if it is neutral. Fluid mechanics focuses on the mathematical description of gas and liquid state of matter and can be used to describe plasma as well. This approximation of plasma as a fluid forms the basis of magnetohydrodynamics (MHD) (Sect. 3.12).

The second volume of the *Atmospheric and Space Sciences* will dwell on the basic physics of plasma environments, focusing on Earth's atmosphere-ionosphere and geospace environment and also to some extent on Mars' ionosphere (Sect. 4.9, but the discussed basic theories and topics are insightful in the context of other solar system planets and can help better understand other planetary ionospheres. This volume is thus complementary to the first volume, which focused on the neutral atmospheric processes. Overall, I will specifically focus on planetary and terrestrial ionospheres and thermospheres. In this chapter, I will provide an introduction to space plasma environments (Sect. 1.2), discuss some important cornerstones of the history of space physics (Sect. 1.3), and will finally summarize some basic plasma parameters (Sect. 1.4), some of which will be further used and discussed in the rest of the book.

The structure of the rest of the monograph is as follows: Next chapter will discuss some basics of the electromagnetic theory; Chap. 3 will provide a basic introduction to transport processes; Chap. 4 will present some fundamental features of planetary ionospheres, focusing on the terrestrial and Martian ionospheres; and Chap. 5 will review some basic principles of the dynamics of the atmosphere-ionosphere system and some research highlights, such as meteorological influences on the upper atmosphere and space weather.

1.2 Space Plasma Environments

The vast majority of the matter in the universe ($\sim$99%) is in the plasma state, i.e., the fourth state of matter. Without the help of advanced observational techniques, such as, satellites or telescopes, this knowledge could not have been revealed until the modern age. Obviously, this fact is not something that the natural philosophers could have predicted easily. The Greek philosophers were highly interested in processes taking place in nature, which is why they are often called natural philosophers. One has to realize that their sole motivation was the endless desire to better understand physical processes occurring around themselves without a specific reason. One of the

most common historical misconceptions was that Earth had a very central role in the "observed universe", what we could summarize as the geocentric world view based on the Ptolemaic model[1] of the solar system in which planets orbited on circles (or deferents) with Earth near the center of the deferent.[2] Later this model was replaced by the heliocentric model, which literally means Sun as the center, officially adopted by Copernicus[3] from the Greek philosophers Hereklid and Samos, who lived 3–4 centuries B.C., and applied in an astronomical context. According to the Copernican revolution, the planets move in an orbit around Sun and the orbit can be approximated by a circular motion.

There is much more to the universe than our planet Earth and the Solar System in which Earth resides as a seemingly extraordinary terrestrial planet. Why is Earth so extraordinary? It is special because many millions of complex living species are thriving on Earth. Mankind has so far not discovered any sign of life in the rest of the universe outside Earth, not even a primitive single cell organism. Despite the fact that we have increasingly better characterized other solar system planets as a result of sophisticated planetary science missions (e.g., MESSENGER, Venera[4] missions, Venus Express,[5] MAVEN,[6] ExoMars,[7] Juno[8], Cassini, Voyager 1 and 2, New Horizon), Earth has kept its special place as a spectacularly habitable planet that happens to have just the optimal conditions for life to develop and be sustained over large period of time. Although science can identify some of the paleontological processes that enables the formation of life, there is not a scientifically convincing argument why Earth happens to have such ideal conditions for life and why the other Solar System planets are so hostile to life. Here we define life with respect to the organic processes that are familiar to us.

From an objective point of view, you just need to check out how the atmospheric environments of our two neighboring planets, Mars and Venus, probably the best known two planets other than Earth, are configured. Venus' atmosphere is way too hot and have extremely dynamic winds. On the other hand Mars appears to be a deserted planet, despite its beautifully reddish color in observations. Needless to mention that Mars surface pressure is less than one percent of Earth's standard pressure as a reason for many challenges, massive planet circumventing giant dust storms could be a major problem for any large-scale population of the Martian surface. Of course, in order to realize the complexity of our universe, humans first had to expand their horizon and allow their curiosity to drive the search for other "realms". With an

[1] The Greek astronomer Ptolemy.

[2] Dictionary of geophysics, Astrophysics, and Astronomy edited by Richard A. Matzner, 2001.

[3] Nicolaus Copernicus (in German: Nikolaus Kopernikus), Polish mathematician and astronomer, 1473–1543.

[4] Russian for Venus. A series of 16 flyby, orbital, and landed missions by the Soviet Union between 1961 and 1983.

[5] First European Space Agency mission to Venus, November 2005–April 2006.

[6] Mars Atmosphere Volatile EvolutioN, http://mars.nasa.gov/maven/.

[7] http://exploration.esa.int/mars/.

[8] https://www.nasa.gov/mission_pages/juno/

impressing speed, more and more exoplanets are discovered and an entire new field of exoplanet discovery and characterization emerged. Currently more than 2000 planets have been discovered, including a very diverse spectrum of planets. Most likely these exoplanets have a spectum of very exotic planetary atmospheres and ionospheres. Better understanding solar system planets can help understand future discoveries of complex structures in exoplanetary atmospheres.

So, where does plasma occur? In small scales for example, the human blood contains a significant portion of non-cellular structures that is referred to as the blood plasma, which contains electrolytes (charged particles). In much larger scales, plasmas are found in the geospace environment and in the far sides of the universe. The most known plasma environments are stellar interiors and atmospheres, stellar winds, planetary upper atmospheres, gaseous nebulae, pulsars, and interstellar hydrogen (H). In the solar system, plasma can be found everywhere and the physics of plasmas is studied in all space weather environments—Sun, heliosphere, magnetosphere, and ionosphere-atmosphere, which I had introduced in the first volume of my book (Yiğit 2015). Closer to Earth, the Van Allen radiation belt, the solar wind and the ionosphere are the most investigated plasma environments. *Van Allen* radiation belt is the region where the charged particles are trapped by the geomagnetic field. Electric and magnetic fields greatly influence the plasma (Chap. 2). The solar wind originating at Sun and propagating with several hundreds of kilometers per second toward Earth impinges on the magnetosphere, which shields the atmosphere and life from dangerous radiation, is significantly distorted in this process. The solar wind consists mainly of electrons and protons and it possesses high conductivity. The typical values of the electron density and temperature in the solar wind are $n_e = 5\,\mathrm{cm}^{-3}$ and $T_e \approx 10^5\,\mathrm{K}$.

Depending on the intensity of the short-term activity on Sun and the configuration of the interplanetary magnetic field (IMF), i.e., the extension of Sun's magnetic field environment into the heliosphere, energy can be transferred from the magnetosphere to the ionosphere to varying extent. Under favorable interplanetary magnetic field conditions and strength, geomagnetic storms can occur. The impact of magnetic storms on the ionosphere and the upper atmosphere are an active subject of research (Yiğit et al. 2016) as will be discussed to some extent in Chaps. 4 and 5.

Figure 1.1 shows the Sun-Earth system and how the collection of the energetic particles originating at Sun, the solar wind, impact the magnetosphere and ultimately Earth's atmosphere. The dayside magnetosphere is compressed while the nightside expands into a tail structure. The IMAGE (Imager for Magnetopause-to-Aurora Global Exploration) and Cluster spacecrafts are situated within the magnetosphere, as marked in the figure. The ionosphere is the weakly ionized portion of the upper atmosphere extending from $\sim$60 km upward and coexists with the thermosphere in the upper atmosphere. The ionosphere is coupled to the thermosphere via composition, dynamics, electrodynamics, and thermal processes. Internal waves generated in the lower atmosphere can propagate upward and greatly influence the thermosphere-ionosphere (Fritts and Lund 2011; Yiğit and Medvedev 2015). This dynamical coupling between the lower atmosphere and upper atmosphere in form of exchange of momentum and energy is called *vertical coupling*.

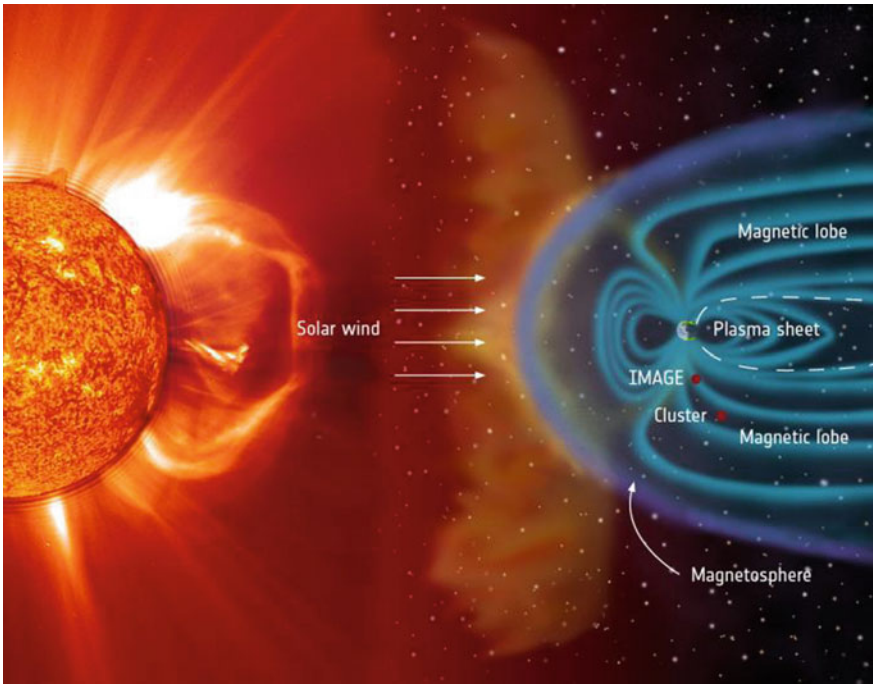

Fig. 1.1 Illustration of how Sun's solar wind affects Earth's magnetic field environment. Within the magnetosphere, the ESA/NASA Cluster and NASA's IMAGE (Imager for Magnetopause-to-Aurora Global Exploration) spacecrafts (*red* markings) survey magnetospheric changes and determine how the solar wind affects the magnetic environment in the geospace (Credit ESA/NASA/SOHO/LASCO/ EIT)

If we consider only the lower atmosphere where our daily lives take place, it is not apparent that the vast majority of the universe is in the plasma state; one would only occasionally encounter plasma in our environment. When there is a severe weather event, localized thunderstorms may take place with intense ionization.

The Northern lights (Aurora Borealis) seen in beautiful colors in Earth's atmosphere are an example of plasma phenomenon. They are generated by the collision of atmospheric neutral gas particles with the charged particles of solar origin entering the atmosphere. Various atmospheric species are thus excited to a higher energy level by particle impact and produce optical emissions upon returning back to the initial state (i.e., ground state). For example, the green aurora is generated due to the collisions of oxygen molecules with solar energetic particles. Figure 1.2 shows how the green aurora looks like from the International Space Station that orbits Earth at about 400 km. Depending on the quantum mechanical energy level transitions within a gas species, other colors of aurora are possible in the Solar System. Earth is not the only planet where auroral activity exists. Despite their large distance

Fig. 1.2 *Green* aurora borealis (or northern lights) over Quebec, Canada, captured in Feb 2012 by an astronaut aboard the International Space Station. The fainter arc of light parallel to the horizon is from airglow (Credit: NASA)

from Sun in comparison with Earth, significant amount of auroral emissions can occur on the giant planets, Jupiter and Saturn, due to the solar wind (Clarke et al. 2009) in additional to strong internal processes.

If the solar radiation provides the initial excitation energy, then the subsequent emissions are called airglow (or sometimes called nightglow). It is a naturally created optical phenomenon by the atmosphere and has a layered structure as can be seen in the background of the aurora in Fig. 1.2. The mesospheres and lower thermosphere (MLT) contains a rich spectrum of emissions due to atomic and molecular oxygen, and sodium (green, blue, and yellow airglow). Red airglow due to atomic oxygen transitions in the upper thermosphere is also seen in the figure.

Next, I will briefly discuss some relevant historical developments in space plasma research.

1.3 A Brief History of Space Research

So far we have discussed what a plasma is and that magnetism and electric forces greatly influence the state and evolution of plasma. Magnetic and electric effects are two interrelated topics that are central to space physics. Magnetism is a consequence of the motion of moving charges, for example, electrons, while electric effects are independent of motion and influence a given charged particle whether it is moving or not. Some selected historical developments concerning magnetism, space plasma research, space missions that we discuss in this section are summarized in Table 1.1.

Table 1.1 Some selected key events in planetary science, space plasma physics, and magnetism from the detection from (1600s) to Solar System Exploration missions of 20th and 21st century

Date	Event
4–5 cc B.C.	Democritus' concept of atom
1600s	Detection of the geomagnetic field
1844	11-year solar cycle proposed by Schwabe
1865	Maxwell's equations
1879	Characterization of aurora by Bequerel
1880	Suggestion of feasibility of achieving orbital velocities by rocket and of satellites (Tsiolkovsky)
1897	Discovery of electron by Thomson
1901	Transatlantic radio transmission by Marconi
1911	Milikan measures the charge on the electron
1919	Rutherford discovers the proton
1929	Introduction of the term "ionosphere" by Watson-Watt
1930s	Interpretation of sunspots as indicators of solar activity by Chapman and Ferraro
1957	Sputnik I and Sputnik II launched by Russians
1958	Detection of radiation belt by Explorer 1, the first US satellite
1961	Soviet Union's Vostok 1 spacecraft takes the first human, Yuri Gagarin, into space
1962	Soviet Union launches Mars-1
1965	NASA's Mariner 4 mission performs the first successful fly-by of Mars
1997	Cassini mission (ESA) launch
Dec 2003	ESA's Mars Express spacecraft enters orbit
2006–2015	New Horizons mission (NASA) launch and arrival at Pluto
Nov 2014	NASA's MAVEN spacecraft enters orbit around Mars
Jul 2016	NASA's Juno spacecraft arrives at Jupiter
Oct 2016	ESA's ExoMars orbiter arrives at Mars

Since Maxwell has unified the electric and magnetic phenomena in 1865 via his infamous equations called *Maxwell's equations* (see Sect. 2.6), we speak of electromagnetism. A better understanding of electromagnetic phenomena led to rapid developments in the exploration of the Solar System.

1.3.1 From the Structure of Matter to the Solar System

Without a deeper understanding of the microscopic structure of matter, plasma could not be better understood. Already during the early times of natural philosophers it has been speculated that the matter in the universe consists of some fundamental

microscopic components, which they called the atom. In fact, it was Democritus[9] who assumed for the first time that everything in nature must be made out of invisible small particles, which he called "atom". Although, Democritus' idea of atom is revolutionary in the ancient times, in the late 19th century, scientists have discovered that the atom is not the smallest indivisible component of matter. Although, the atom is the smallest structure of a chemical element, which by definition consists of atoms that have nuclei with the same charge, there are smaller fundamental structures that make up the atom, i.e., electrons, protons, and neutrons, which themselves are made out of various flavors of quarks. One of the earliest models of the atom was introduced by Thomson who discovered in 1897 the electron and determined experimentally the electron charge-electron mass ratio, e/m_e. He won the Nobel Price in 1906 owing to this discovery. Using the electron charge-mass ratio Milikan determined in 1911 the charge on the electron. Proton was discovered later. Bombarding nitrogen nuclei with energetic α-particles from a naturally radioactive source, Rutherford demonstrated in 1919 that hydrogen nuclei were produced as a result of this interaction. The hydrogen nucleus was called proton at that time.

Magnetism plays a key role in the history of space research. Already more than two thousand years ago the ancient Greeks have discovered magnetism after they observed that certain stones attract iron. Our planet Earth demonstrates fascinating magnetic properties. From the point of magnetism, to a first approximation, it can be viewed as a giant dipole magnet, meaning that it has two magnetic poles. The geomagnetic field B_E itself has not been detected until 1600. William Gilbert suggested in his work *De Magnete* that the north-south alignment of the compass results from the geomagnetic field, stating for the first time that Earth is a giant magnet. Carl Friedrich Gauss (1777–1855) suggested that the geomagnetic field is composed of an internal field and of a component that is generated in the atmosphere. Gauss and Wilhelm E. Weber (1804–1891) initiated precise measurements of the geomagnetic field. In the 19th century global network of observatories have increasingly been established.

Magnetic activity and solar activity are two connected processes. In 1844, Schwabe proposed an approximate ten-year solar cycle variations as reported in the German publication "Sonnen-Beobachtungen im Jahre 1843" (Solar observations in the year 1843) in which he presented sunspot group numbers from 1829 to 1843 (Schwabe 1844). Sir Edward Sabine proposed the 11-year periodicity in the magnetic cycle according global network of measurements.

The modern era of rocket propulsion has been launched with the ideas of the Russian and Soviet scientist Konstantin Tsiolkovsky (1857–1935) in 1880. An image of this revolutionary scientist can be found in the Tsiolkovsky Museum in Russia (Fig. 1.3). Tsiolkovsky was the first who has discussed the feasibility of achieving orbital velocities by rockets and weightlessness and predicted the development of modern-era satellites. He has discussed his ideas in his famous book "Dreams of

[9]In German: Democritus; an ancient Greek pre-socratic natural philosopher who was born in 460 B.C. in Thrace, which is situated currently in Turkey and is called "Trakya", that is the European portion of Turkey. he died in 370 B.C.

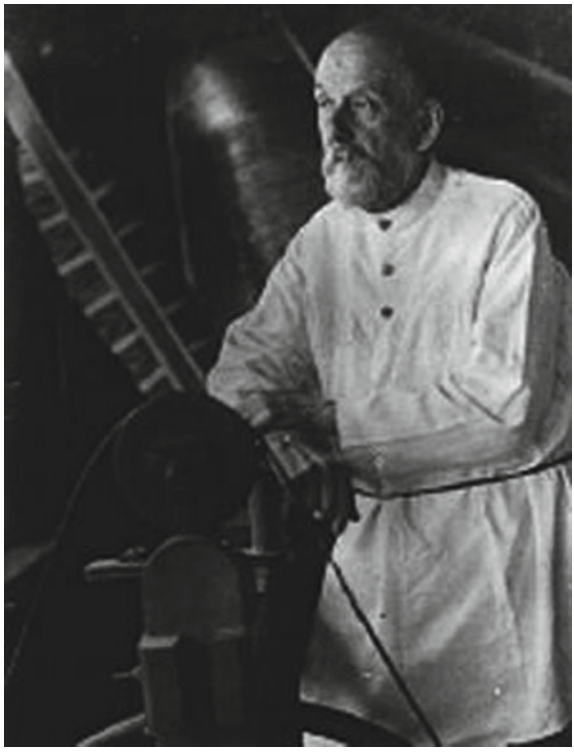

Fig. 1.3 An image of Konstantin Tsiolkovsky. He is considered one of the leaders of original thinkers in the history of spaceflight Image Credit: K. E. Tsiolkovsky Museum, Kaluga, Russia/NASA

Earth and Sky", originally published in 1895. He even studied the possibilities of constructing a space station that would orbit Earth. Owing to his immense creativity for his time, he is considered one of the leaders of original thinkers in the history of spaceflight as he used the concept of "thought experiments" that were often very efficiently used by prominent physicists familiar to the general public (Pearson 1997). A crater on Moon, the Tsiolkovsky crater, has been named after him. It was not until 1957 that the first artificial satellite, Sputnik 1, was launched into an elliptical low orbit by the Soviet Union. As an appreciation of his visionary ideas, the Sputnik 1 launch was made by the Soviets to coincide with Tsiolkovsky's centenniel. This satellite stayed in the orbit for 92 days, having demonstrated clearly the feasibility of satellite technology. Today, hundreds of satellites are orbiting Earth.

The solar activity has been quantified since 1610s in terms of the sunspots, which are magnetically dense, dark, and cold areas of the solar photosphere (Clette et al. 2014). The relationship between sunspot numbers and the frequency of geomagnetic disturbances have been discovered in 1952. In 1867, Angstrom has performed the first observations of the auroral spectrum (Angstrom 1867). Late 1870s Henri Bequerel stated that the sunspots are sources of fast protons and that as these particles reach Earth, they are guided toward the auroral oval by the geomagnetic field (Bequerel 1879). In his book "Aurorae: their characters and spectra" published in 1879, he has

provided a detailed discussion of the general and specific characters of the aurora, including observations by the ancient Greeks.

In the early 20th century, Birkeland has provided an experimental evidence for the aurora in a laboratory simulation called the terrella. His tests predicted that the geomagnetic field is responsible for the auroral oval. In 1901 Marconi has performed the first transatlantic radio transmission, which eventually led to the development of the physical theory of the ionosphere starting with the work of Lodge (1902) who has pointed out that the solar radiation produces a conducting layer which reflects radio waves. Though, it is interesting to note that in 1830s Gauss had speculated on the existence of a conduction layer in the atmosphere. In 1920s, the existence of the ionosphere has formally been established. As noted in the seminal historical review by Rishbeth (2001), the first formal mention of the term "ionosphere" was apparently in the work by Watson-Watt (1929).

In 1930s, Sydney Chapman and Victor Ferraro suggested that sunspots are indicators of solar activity and that this activity is related to solar flares, which are giant eruptions in the atmosphere of Sun. In situ measurements have demonstrated the complexity of the geomagnetic field. The second half of the 20th century marked the application of satellite technology in surveying, in particular, the upper atmosphere. Despite the much rarefied nature of the upper atmosphere, its technological significance has been recognized. The upper atmosphere is the region where satellite missions take place and spacecraft enter an orbit. If a satellite changes its orbit and enters a denser atmospheric region, the spacecraft may experience a larger drag, which then has impact on its lifetime.

1.3.2 Post-World War-II Period

The second half of the 20th century deserves special attention in the history of space sciences. As soon as the Second World War ended in 1945, the humanity has become increasingly curious about how to exploit the technologies that have been designed or even fully developed during the war for better purposes. Russians have launched Sputnik I on 4th October 1957, which marked the official beginning of the space age. The Sputnik spacecraft was the first man-made object ever to leave Earth's atmosphere. It was a mere 83 kg satellite that stayed in the orbit for about three months. Sputnik 2 was launched about a month later (3 November 1957) and actually carried the first living being, a dog called Leika, into an orbit around Earth. The Sputnik program had an immense political impact on the world politics as well because it was used by some ideological groups as a symbol of the superiority of the Marxist-Leninist Soviet Union over the predominantly capitalist Western powers.

In 1958, the first US satellite, Explorer 1, detected Earth's radiation belt using a Geiger counter (Allen et al. 1958). The Explorer 1 team, including William H. Pickering, former director of JPL (left), James A. van Allen (center), of the State University of Iowa, who designed and built the instrument on Explorer, and Wernher von Braun (right) are seen in Fig. 1.4 holding the Explorer 1. In 1961, Soviet

Fig. 1.4 Explorer 1 rocket and the team members William H. Pickering (*left*), James A. van Allen (*center*) of the State University of Iowa, and Wernher von Braun (*right*), leader of the Army's Redstone Arsenal team Credit: NASA

Union's "Vostok program" was the first to put the first human into space, in Vostok 1 spacecraft. Since 1960s, many satellite missions have provided a more detailed understanding of the atmosphere as a whole.

Overall, technological advancements together with a better understanding of the universe have even enabled the development of satellite missions to other planets in the Solar System. Mars 1 launched by the Soviet Union in November 1962 was one of the first mission that has approached Mars.

One of the most successful ones is the Cassini mission, which was started by the European Space Agency (ESA) in 1997. Due to its success, this mission has been extended as Cassini Equinox and more recently as Cassini Solstice mission.[10] Of course, Cassini is not the most recent space mission. ESA with the successful participation of several international agencies placed the Mars Express spacecraft into orbit around Mars in December 2003. The orbiter is still operational and have been invaluable for the planetary science community. Such an orbiter inhabits several instruments, which are capable of observing the environment of a planet. The collection of the different instruments onboard the spacecraft is called a payload.

[10] http://saturn.jpl.nasa.gov/.

Table 1.2 Mars Express payload. Mars Express entered orbit around Mars in Dec 2003

Instrument	
HRSC	High Resolution Stereo Camera
OMEGA	Visible and Infrared Mineralogical Mapping Spectrometer
MARSIS	Sub-surface Sounding Radar Altimeter
PFS	Planetary Fourier Spectrometer
SPICAM	Ultraviolet and Infrared Atmospheric Spectrometer
ASPERA	Energetic Neutral Atoms Analyser
MaRS	Mars Radio Science Experiment

Fig. 1.5 An image captured in January 2017 by the HRSC instrument onboard Mars Express during Mars Express Orbit 16565 showing southern highlands of Mars where an exotic triple crater is found. The image is centered around 198 °E and 27 °S with a resolution of ∼22 m per pixel Credit: ESA/DLR/FU Berlin

For example, the High Resolution Stereo Camera (HRSC) is able to capture very high-resolution images of the Martian surface. Table 1.2 lists the other instruments onboard Mars Express.

Figure 1.5 illustrates the level of sophistication in imaging Mars' surface. Mars Express' HRSC instrument took this image in January 2017 highlighting a triple-crater in Martian southern highlands. It is one of the many images, which demonstrate that the Martian surface has been heavily bombarded by meteorites/meteors of various sizes.

Overall, each planet in the Solar System has been visited by a space probe. Most recently, the Trace Gas Orbiter (TGO) of ESA's ExoMars mission has successfully

arrived at Mars in October 2016 and entered an elliptical orbit around the "Red Planet".

Even Pluto, currently accepted as a dwarf planet, has recently been visited by NASA's New Horizon mission,[11] which is revealing amazing features about this exotic dwarf planet. Also, there are a number of planned missions to explore the Solar System. Overall, it is just a matter of time that the humanity will be able to settle down in new worlds.

1.4 Plasma Parameters and Characteristics

Two important questions about plasma properties are how hot and how dense the plasma is. The temperature and density of the plasma in the universe can be characterized by the electron temperature T_e and electron number density n_e. The temperature is a measure of the thermal energy or in other words the average kinetic energy $\overline{E}_k = 3/2\, kT_e$, where the temperature is often specified in terms of units of particle energy, i.e., electron volts with

$$1\ \text{eV} = 1.6022 \times 10^{-19}\ \text{J} \tag{1.3}$$

and

$$k = 1.38 \times 10^{-23}\,\text{J}\,\text{K}^{-1} \tag{1.4}$$

is the Boltzmann constant.[12] Note how large one Joule is in terms of one electron volt. One electron volt is the energy a particle carrying a charge of e gains or losses in falling a potential drop/difference of 1 V (volt), where the potential difference is the line integral of the electric field strength between the reference levels:

$$\Delta V = V_1 - V_2 = \int_1^2 \text{E} \cdot \text{ds}, \tag{1.5}$$

which expresses an energy difference of 1 J per unit charge.

Using the magnitude of $e = 1.6022 \times 10^{-19}$ C to define a unit of energy serves a convenient measure describing atomic systems and various aeronomic phenomena. For example, electron and proton precipitations in the ionosphere, in general, termed as particle precipitation, are extremely variable and has typical energies of keV (10^3 V) to MeV (10^6 V) as derived from the Defense Meteorological Satellite Program (DMSP) satellite observations (Galand et al. 2001). Typically, the plasma temperature and density are given in terms of T_e and n_e, i.e., the number of electrons N_e per unit volume. While, the electrons are the mobile part of a plasma, the much heavier

[11] https://www.nasa.gov/mission_pages/newhorizons/main/index.html.

[12] Named after the German physicist L. Boltzmann (1844–1906).

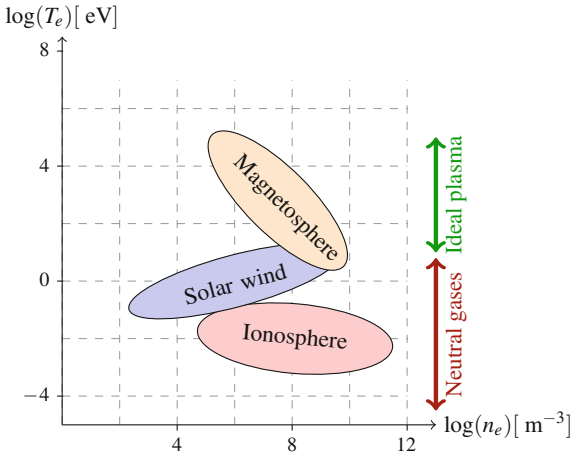

Fig. 1.6 Three regions, magnetosphere, solar, wind, and ionosphere, where astrophysical plasmas are found, are illustrated in a $T_e - n_e$ diagram. Plasma temperature T_e is in eV and the plasma density n_e is in m^{-3}

ions are rather immobile and are thus regarded during many collisions as a fixed background of positively charged particles. If n_e is relatively large/small, then one speaks of a high-density/low-density plasma. The morphology of the ambient electric and magnetic fields affects the plasma motion.

The electron temperature and number density classify with respect to

1. interactions between plasma constituents,
2. relative importance of **E** and **B** fields for the particle motion,
3. range over which particles can propagate freely.

Plotting the plasma temperature as a function of density as $\log(T_e) - \log(n_e)$ diagram for astrophysical plasmas helps categorize various plasmas, as illustrated in Fig. 1.6. Typically, at low temperatures, the neutral gases are partially ionized, such as the case in the terrestrial ionosphere and other planetary ionospheres in the solar system. The peak electron number density found around 300 km in Earth's ionosphere is in the order of 10^{12} m^{-3}, where the background thermosphere has a density of 10^{14} m^{-3}, suggesting that the neutral atmosphere is about 100 times denser than the terrestrial ionosphere. With increasing temperature one obtains an ideal plasma, which is fully ionized, as it is the case in the magnetosphere. The necessity of very high temperatures is the reason why we do not encounter ideal plasma on Earth under natural conditions.

Another important physical process in plasmas is *Debye shielding*. It is the shielding of a given charged particle in a plasma by a cloud of oppositely charged particles such that the Culomb field of an individual charged particle is not felt. The associated length scale is known as the *Debye length* λ_D given by

$$\lambda_D = \left(\frac{\varepsilon_0 k T_e}{e^2 n_e} \right)^{1/2}, \tag{1.6}$$

where

$$\varepsilon_0 = 8.854 \times 10^{-12}\ \mathrm{F\,m^{-1}} \tag{1.7}$$

is the permittivity of free space (or vacuum) and T_e is the electron temperature, In Earth's F region ionosphere with $n_e \approx 10^{12}\ \mathrm{m^{-3}}$ and $T_e \approx 1500$ K, this scale is in the order of millimeters (10^{-3} m). This length scale describes the spatial scale over which particles in a plasma exert electrostatic forces on each other. In other words, despite the long-range effect of Culomb forces, charged particles in a plasma do not interact at a distance greater than the Debye length. Consider a certain number N_D of particles inside an imaginary sphere, i.e., a Debye sphere, of radius λ_D. Then, the Debye length is inversely proportional to the number of particles, $\lambda_D \propto N_D^{-1}$, because in a dense plasma, charges of opposite signs screen each other, so that the length over which the electrostatic forces are exerted is smaller, meaning a stronger Debye shielding. The Debye length is proportional to the plasma temperature and inversely proportional to the plasma density, $\lambda_D \propto \sqrt{T_e/n_e}$. In a hotter plasma, thermal motion increases and the quasi-neutrality can be violated over a larger scale.

Other important parameters are the ion temperature T_i, gyrofrequency (or cyclotron frequency), gyroradius (or cyclotron radius) r_c, and the plasma frequency ω_{p_s} (of species s). The gyrofrequency, as the name suggests, is the frequency with which charged particles gyrate around magnetic fields. It is proportional to the charge and the magnetic field strength and inversely proportional to the mass of the plasma species

$$\omega_{c_s} = \frac{|e|B}{m_s}. \tag{1.8}$$

The radius associated with the charge particle gyration orbit is called the gyroradius (also called cyclotron radius or even Larmor radius) and obtained from the balance the magnetic force and the centripetal force in the particle orbit:

$$|e|v_\perp B = m\frac{v_\perp^2}{r_c}, \tag{1.9}$$

which gives

$$r_c = \frac{mv_\perp}{|e|B}, \tag{1.10}$$

where $v_\perp$ denotes the component of the instantaneous charged particle velocity perpendicular to the magnetic field. Additionally, the gyrofrequency and the gyroradius are related by

$$\omega_c = \frac{v_\perp}{r_c}. \tag{1.11}$$

In a plasma positively charged ions are an important species. As they contain much heavier protons and neutrons, ions are much heavier than the electrons, $m_i > m_e$.

Thus, for a given ambient magnetic field, electrons gyrate faster around the magnetic field than ions, $\omega_{c_e} > \omega_{c_i}$. For example, for an electron in an ambient magnetic field strength of 20 000 nT at F-region altitudes we get a gyrofrequency of $\omega_{c_e} = |e|B/m_e = 20 \times 10^{-6}\,\text{T} \times 1.602 \times 10^{-19}\,\text{C}/(9.109 \times 10^{-31}\,\text{kg}) \approx 3.52 \times 10^6\,\text{Hz}$, which is more than one million oscillations per second. For the same conditions, for an atomic oxygen ion we get $\omega_{cO+} \approx 120$ Hz. It is clearly seen that because $\omega_{c_e} \gg \omega_{cO+}$ the lighter electrons are much more confined to the magnetic field than the heavier oxygen ion, which completes only 120 cycles per second, corresponding to a period of 0.05 s.

The plasma frequency is the frequency that describes plasma oscillations that occur, for example, in response to external magnetic fields. If electrons are displayed from the ions, electric fields build up, which accelerate the lighter electrons. Electrons then overshoot due to their inertia and thus start oscillating around an equilibrium position with a characteristic frequency called the *plasma frequency*. During these oscillations the much heavier ions do not have time to respond to the electron motion and can therefore be considered as a fixed background. The electron plasma frequency is given by

$$\omega_{pe} = \left(\frac{n_e e^2}{\varepsilon_0 m_e}\right)^{1/2}. \tag{1.12}$$

Because the electron mass is very small, ω_{pe} is a large frequency in the order of MHz, corresponding to the scale of radio wave frequencies.

Table 1.3 summarizes some important plasma parameters I will be discussing in this book. The next chapter will introduce some basic principles of electromagnetism that is relevant to the context of this book.

Table 1.3 Some common plasma parameters and fields. The subsript s denotes a given plasma species, e.g., electrons or ions. Sometimes gyroradius is referred to as Larmor radius

Notation	Plasma parameter
n_e	Electron number density
T_e	Electron temperature
T_i	Ion temperature
ω_{c_s}	Plasma cyclotron frequency/gyrofrequency (of species s)
r_{c_s}	Plasma cyclotron radius/gyroradius (of species s)
ω_{p_s}	Plasma frequency (of species s)
λ_D	Debye length
E	Electric field strength
B	Magnetic field

References

Allen JAV, Ludwig GH, Ray EC, McIlwain CE (1958) Observation of high intensity radiation by satellites 1958 alpha and gamma. J Jet Propuls 28(9):588–592

Angstrom M (1867) Observations on certain lines of the solar spectrum. Philos Mag Ser 4(33):76–78

Bequerel H (1879) Aurorae: their characters and spectra. Spon, London

Clarke JT, et al (2009) Response of Jupiter's and Saturn's auroral activity to the solar wind. J Geophys Res 114:A05210. doi:10.1029/2009JA013694

Clette F, Svalgaard L, Vaquero JM, Cliver EW (2014) Revisiting the sunspot number. a 400-year perspective on the solar cycle. Space Sci Rev 186:35–103

Fritts DC, Lund TC (2011) Gravity wave influences in the thermosphere and ionosphere: observations and recent modeling. In: Aeronomy of the Earth's atmosphere and ionosphere, IAGA special sopron book series. Springer, Netherlands, pp 109–130. doi:10.1007/978-94-007-0326-1

Galand M, Fuller-Rowell TJ, Codrescu MV (2001) Response of the upper atmosphere auroral protons. J Geophys Res 106(A1):127–139

Lodge O (1902) Mr. Marconi's results in day and night wireless telegraphy. Nature 66(199):222

Pearson J (1997) Konstantin Tsiolkovsky and the origin of the space elevator. In: 48th international astronautical congress, Turin, Italy, pp IAF–97–IAA.2.1.09

Rishbeth H (2001) The centenary of solar-terrestrial physics. J Atmos Sol-Terr Phys 63:1883–1890

Schwabe H (1844) Sonnen-beobachtungen im Jahre 1843. Astronomische Nachrichten 21:233–234

Vasyliunas VM (2001) Electric field and plasma flow: what drives what? Geophys Res Lett 28(13):2177–2180

Vasyliunas VM (2010) Physics of magnetospheric variability. Space Sci Rev. doi:10.1007/s11214-010-9696-1

Watson-Watt RA (1929) Weather and wireless. Q J R Meteorol Soc 55:273–301

Yiğit E (2015) Atmospheric and space sciences: neutral atmospheres, vol 1. Springer briefs in earth science. Springer, Netherlands. doi:10.1007/978-3-319-21581-5

Yiğit E, Medvedev AS (2015) Internal wave coupling processes in Earth's atmosphere. Adv Space Res 55(5):983–1003. doi:10.1016/j.asr.2014.11.020, http://www.sciencedirect.com/science/article/pii/S0273117714007236

Yiğit E, Knížová PK, Georgieva K, Ward W (2016) A review of vertical coupling in the atmosphere-ionosphere system: effects of waves, sudden stratospheric warmings, space weather, and of solar activity. J Atmos Sol-Terr Phys 141:1–12. doi:10.1016/j.jastp.2016.02.011, http://www.sciencedirect.com/science/article/pii/S1364682616300426

Chapter 2
Basic Electromagnetic Theory

Electromagnetism

The man who does not read good books has no advantage over the man who cannot read them.

Mark Twain (1835–1910)

Abstract Physics of plasmas is described by the electromagnetic theory. Elementary principles of electromagnetism are reviewed in this chapter. Electric fields, magnetic fields, and Maxwell equations, which self-consistently relate these fields to each other, are introduced. Charged particles are sources of electric fields and moving charged particles can generate magnetic fields. The most obvious manifestation of electromagnetic fields is the light. The elecromagnetic spectrum is composed of a broad range of wavelengths from low-frequency radio waves ($\sim$10 MHz) to very high-frequency gamma rays ($\sim 10^{27}$ Hz). Finally the concept of phase space and collisions between particles of different species are discussed.

Keywords Electric field · Magnetic field · Electromagnetism · Magnetic flux · Electric currents · Maxwell equations · Poynting flux · Phase space · Collisions · Summation convention

2.1 Electric Field

Around a region of some electric charge distribution σ, there exists an electrostatic field and thus an electric field **E**. The simplest scenarios of a positively and a negatively charged particles are illustrated in Fig. 2.1. The term "electric" is of Greek origin, "elektron", and means amber. If there are any other charged particles within this field, that is in the space around σ, they experience a force, which is called the electric force. Strictly speaking, this force is said to be of electric origin if and only if it exists when a charged particle is inserted into the field. The resulting electric force on the inserted test particle at a point is proportional to the strength of the electric field and the charge intensity:

E. Yiğit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*, SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-62006-0_2

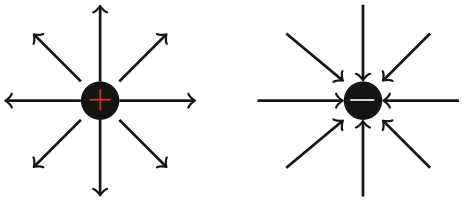

Fig. 2.1 Illustration of the electric field around a positive (*left*) and a negative charge (*right*). For a positive point charge, the field lines diverge, while they converge for a negative point charge

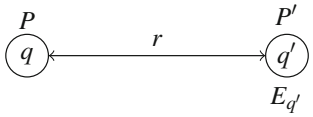

Fig. 2.2 The electric field $|\mathbf{E}| = E_{q'}$ at point P' experienced by a test charge q', distant r away from a point charge q located at point P

$$|\mathbf{F}_E| = |\mathbf{E}|q. \tag{2.1}$$

The electric field $\mathbf{E} = (E_x, E_y, E_z)$ is a vector field and the direction of the force is given by the direction of the electric force on a positive charge. The unit of the electric field is Newton per Coulomb (N C^{-1}). It is instructive to ask how one can determine the electric field distribution around a charged particle. To find the electric field at point P, distant $r = r_{qq'}$ away from a point charge q, imagine a test charge q' at P', as illustrated in Fig. 2.2. The force between the two charges is given by the *Coulomb's law*:

$$F = \frac{1}{4\pi\varepsilon_0}\frac{qq'}{r^2}, \tag{2.2}$$

where

$$\varepsilon_0 = \frac{1}{(\mu_0 c^2)} = 8.854 \times 10^{-12}\ \text{F m}^{-1} \tag{2.3}$$

is the permittivity of free space and

$$\frac{1}{4\pi\varepsilon_0} \approx 9 \times 10^9\ \text{m F}^{-1}. \tag{2.4}$$

Then, the electric field $E_{q'}$ at the position of the test charge q' is

$$E_{q'} = \frac{F}{q'} = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}. \tag{2.5}$$

This formalism can be extended to multiple charges, $q_1, q_2, q_3, \ldots$, with distances $r_{12}, r_{13}, \ldots$, etc. In general, if there are N number of point charges, the resulting

electric field at the location P' of the test charge is a superposition of all the individual ambient electric fields associated with other point charges:

$$E = \frac{1}{4\pi\varepsilon_0}\sum_{i}^{N}\frac{q_i}{r_i^2}. \tag{2.6}$$

2.2 Elecric Potential Energy

As an electric field exerts a force on a charged particle in the vicinity, work is performed, given by

$$W_{1\to 2} = \int_1^2 F_E\, ds \tag{2.7}$$

in moving the charge from location 1 to location 2. This work is independent of the path taken between the initial and the final locations and is thus said to be "conservative", just like the force performed in a gravitational field. This work can be represented by a potential energy E_p

$$E_p = \int_1^2 E\, q\, ds. \tag{2.8}$$

Generalizing this concept to a static charge distribution yields:

$$E_p = \frac{1}{4\pi\varepsilon_0}\sum_{i<j}\frac{q_i q_j}{r_{ij}}. \tag{2.9}$$

Using the potential energy, an electric potential V at any point in an electric field can be defined as the potential energy per unit charge

$$V = \frac{E_p}{q}. \tag{2.10}$$

Electric field and potential are closely related. The potential difference between two points "a" and "b" is the line integral of the electric field strength.

$$V_a - V_b = \int_a^b \mathbf{E}\cdot \mathrm{d}\mathbf{s}. \tag{2.11}$$

This potential difference can be interpreted as the energy required to move a unit charge from a to b within the ambient electric field distribution $\mathbf{E}$ as illustrated in Fig. 2.3. This electrostatic field is conservative, thus

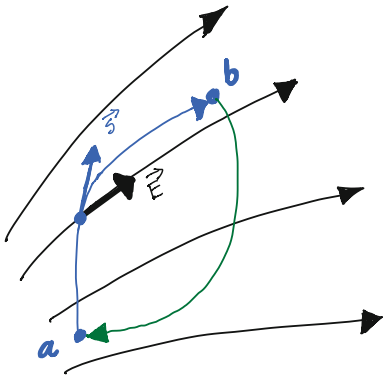

Fig. 2.3 Illustration of the work done in a representative electric field **E**, shown with *black field lines*, by moving a test charge from point a to b (*blue path*) and then from b to a (*green path*), which constitutes a closed path. The orientations of the path vector and the local electric field at a representative point are shown with *thick blue* and *black arrows*, respectively

$$\oint \mathbf{E} \cdot \mathrm{d}\mathbf{s} = 0, \tag{2.12}$$

indicating again that work done in the electric field is independent of the path. So, in a closed path, such as the one illustrated in Fig. 2.3, no net work is performed.

Equivalently, the electric field can be deduced from an electric potential as

$$\mathbf{E} = -\nabla V, \tag{2.13}$$

where the electric field points in the direction opposite from the direction of maximum rate of potential increase. So, as a charge is a sorce of an electric field, it also generates an electric potential field in its vicinity whose strength decreases with increasing distance (i.e., $V \propto q/r$). Note that the electric field strength is inversely proportional to the square of the distance ($V \propto q/r^2$).

2.3 Magnetic Field

The origin of magnetic fields can be understood by considering moving charges. A moving charge sets up in the space around it a magnetic field, which exerts a force on another charge moving through it. So, the magnetic field **B** is said to exist at a point if a force is exerted on a moving charge at that point. The associated magnetic force is

$$\mathbf{F}_B = q\mathbf{v} \times \mathbf{B}, \tag{2.14}$$

or equivalently

$$F_i = q\ \varepsilon_{ijk} v_j B_k, \tag{2.15}$$

where $\mathbf{v} = (v_x, v_y, v_z)$ and $\mathbf{B} = (B_x, B_y, B_z)$ are the three-dimensional charged particle velocity and the magnetic field vector, respectively, and ε_{ijk} is the Levi-Civita

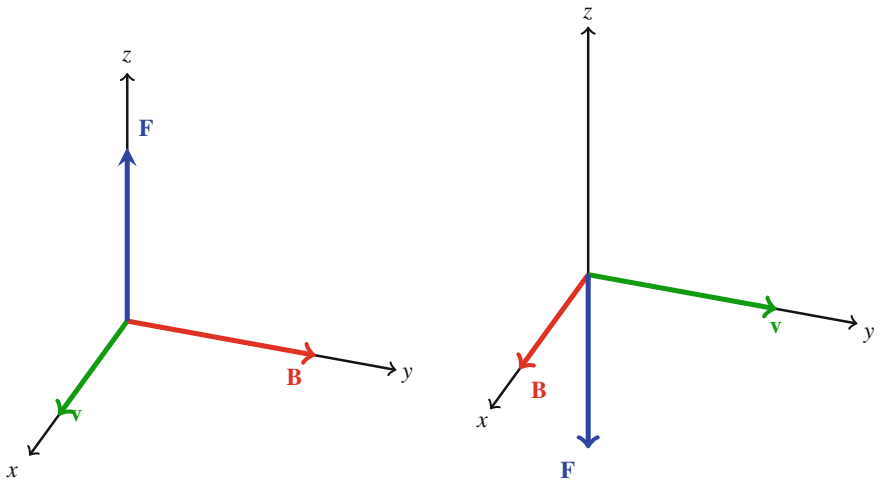

Fig. 2.4 Illustration of the vectors $\mathbf{v} \times \mathbf{B}$ in the rectangular Cartesian coordinate system for two scenarios: charged particle motion is the x-axis and the magnetic field is along the y-axis (*left*); motion is along the y-axis and the field is directed along the x-axis (*right*)

symbol (appendix B.1). The resulting force is perpendicular to both the motion and the field as shown in Fig. 2.4 for two different cases. The cross-product (or vector product) operator "$\times$" implies a counterclockwise rotation described by the rotation angle $\alpha_{\mathbf{vB}}$

$$\mathbf{v} \times \mathbf{B} = v\ B\ \sin\alpha_{\mathbf{vB}} \tag{2.16}$$

where $v = |\mathbf{v}|$, $B = |\mathbf{B}|$, and $\mathbf{v} \times \mathbf{B} = -\mathbf{B} \times \mathbf{v}$, which explains why the force is pointing in the direction of the $-z$-axis in the figure on the right. In practical terms, this cross-product can be represented by the "right-hand-rule", where, e.g., the pointing finger is aligned with the velocity vector, the middle finger shows, with an angle $\alpha_{\mathbf{vB}}$ with respect to the pointing, in the direction of the magnetic field, then the resulting magnetic force is along a line perpendicular to the **v**-**B** plane, shown by the direction of the thumb. The unit of the magnetic field is the tesla (in recognition of Nikola Tesla) and is given by the unit of $F/(qv)$, that is:

$$1\ \text{tesla} = 1\ \text{T} = \frac{\text{N}}{\text{C m s}^{-1}} = \frac{\text{N}}{\text{A m}}. \tag{2.17}$$

Another practical unit of the magnetic field is the gauss G, where $1\ \text{G} = 10^{-4}$ T. The gauss is a useful unit because Earth's magnetic field strength on the surface is in the order of 10 G. In comparison, in a laboratory environment, the strongest field that can be produced is around 5×10^5 G = 50 T.

More precisely, the vector field **B** is called the *magnetic induction field* and the lines of induction describes the magnetic field. The components of the magnetic force field in Cartesian coordinates are given by

$$\mathbf{F}_B = q \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ v_x & v_y & v_z \\ B_x & B_y & B_z \end{vmatrix} = q\,[\,\hat{\mathbf{i}}(v_y B_z - v_z B_y) - \hat{\mathbf{j}}(v_x B_z - v_z B_x) + \hat{\mathbf{k}}(v_x B_y - v_y B_x)], \tag{2.18}$$

where $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are the Cartesian coordinate unit vectors with the properties $\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}, \hat{\mathbf{k}} \times \hat{\mathbf{i}} = \hat{\mathbf{j}}$, and $\hat{\mathbf{j}} \times \hat{\mathbf{k}} = \hat{\mathbf{i}}$.

The field at a point can be characterized by the number of lines per unit area. The line of induction is expressed in the units of Weber (Wb) and the magnetic field is quantified by the unit Wb m^{-2}, which is equal to 10^4 Gauss.

The total number of lines through a surface A defines the *magnetic flux* Φ_m:

$$\Phi_m = \mathbf{B} \cdot \mathbf{n}\, A = B\, A \cos\alpha, \tag{2.19}$$

where $\mathbf{n}$ is the surface normal vector, α is the angle between the surface normal and the magnetic field vector. For a uniform magnetic field that is normal to the surface A, the magnetic flux is $\Phi_m = B\,A$. The magnetic induction is equivalent to the magnetic flux density.

2.4 Electromagnetic Field

The electromagnetic field consists of an electric and magnetic field component. Then, in the presence of an ambient electric and magnetic field, a point charge q moving with velocity $\mathbf{v}$ is subject to an electric force $\mathbf{F}_E$ and a magnetic force $\mathbf{F}_B$:

$$\begin{aligned} \mathbf{F}_L = m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} &= \mathbf{F}_E + \mathbf{F}_B \\ &= q\mathbf{E} + q\mathbf{v} \times \mathbf{B}, \qquad (2.20) \\ &= q(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \qquad (2.21) \end{aligned}$$

where $\mathbf{F}_L$ is called the Lorentz force; electric field and magnetic field in the Cartesian coordinate system are described by the Cartesian unit vectors. In our discussions, the charge speed $v = |\mathbf{v}|$ is assumed to be much smaller than the speed of light

$$c \approx 3 \times 10^8 \text{ m s}^{-1}, \tag{2.22}$$

i.e., ($|\mathbf{v}| \ll c$), so that $\gamma = (1 - v^2/c^2)^{-1/2} \approx 1$ and relativistic effects can therefore be neglected.

The electric field causes an acceleration $(q/m)\mathbf{E}$ that is parallel to the electric field, in the direction of the field for positively charged particles and in the opposite direction of the field for the electrons. The magnetic field causes a centripetal acceleration $(q/m)\mathbf{v} \times \mathbf{B}$ that causes the particle to gyrate in a circular orbit around the magnetic field.

2.5 Electric Currents

Electric currents are caused by the motion of charged particles. The current density $\mathbf{j}$ at a given point in space expresses the amount of electric current per unit area (or simply charge flux). For free electron motion, the associated current density is

$$\mathbf{j}_e = \rho_e \mathbf{v}_e, \tag{2.23}$$

where $\rho_e = |e| n_e$ is the space charge density of electrons. In general, in conducting fluids, the current density is proportional to the electric field:

$$\mathbf{j} = \sigma \mathbf{E}, \tag{2.24}$$

where σ [in Siemens per meter] is the electric conductivity. The relation (2.24) is known as Ohm's law.

Metals are very good conductors ($\sigma \sim 10^7$ S m^{-1}), while fluids have relatively smaller conductivity (sea water: $\sigma \sim 5$ S m^{-1}).

For moving conductors, Ohm's law can be generalized as

$$\mathbf{j} = \sigma (\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{2.25}$$

where all variables are measured with respect to the same fixed frame of reference. The magnetic field is a net field, which includes any internal fields (or seed fields) and the field due to the current density (thus the charge flow). One important manifestation of this principle occurs in sunspots, which are essentially self-excited dynamos (Lorrain and Koutchmy 1998). In sunspots due to the motion of the plasma, a current density $\mathbf{j} = \sigma (\mathbf{v} \times \mathbf{B})$ is induced within the seed field $\mathbf{B}$. The associated magnetic field of the induced current has the same sign as the seed field, thus provides a positive feedback. The resulting magnetic force density [N m^{-3}] is perpendicular to both the current density and the magnetic field:

$$\mathscr{F}_M = \mathbf{j} \times \mathbf{B}. \tag{2.26}$$

In a plasma also ions are mobile and the associated ion current density is

$$\mathbf{j}_i = \rho_i \mathbf{v}_i. \tag{2.27}$$

Therefore, one key plasma property is the net current due to the combined electron and ion motion. Presence of charged particles does not necessarily mean that there exist a net current. If electrons and positive ions move together than no net current is generated, but if they have a relative drift then the resulting current density is

$$\mathbf{j} = e(Z n_i \mathbf{v}_i - n_e \mathbf{v}_e), \tag{2.28}$$

where Z is the number of charges per ion. For example, for O^+ Z = 1. In the case of charge neutrality in a plasma $n_e = n_i$, the current density is

$$\mathbf{j} = en_e(\mathbf{v}_i - \mathbf{v}_e). \tag{2.29}$$

2.6 An Overview of Maxwell Equations of Electromagnetic Field in Vacuum

The unification of the electric and the magnetic fields have been proposed for the first time in 1873 by Maxwell as the electromagnetic (EM) theory. In general, sources of electromagnetic fields are charges, magnetized bodies, and currents. The EM field can be discrete or continuous; and stationary or time-dependent. Maxwell formulated four fundamental equations of the EM field in vacuum:

1. The **Gauss' law** (in differential form) for the electric field states that the electric field is generated by a charge density ρ_c

$$\nabla \cdot \mathbf{E} = \frac{\rho_c}{\varepsilon_0}, \tag{2.30}$$

which can be written in terms of the Poisson's equation as

$$\nabla^2 V = -\rho_c/\varepsilon_0. \tag{2.31}$$

2. **No magnetic monopoles** (charges) or no magnetic currents exist, described by

$$\nabla \cdot \mathbf{B} = 0 \tag{2.32}$$

3. **Faraday's law** states that an electric field is generated by a changing magnetic field:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.33}$$

4. The **Ampere's law** says that a (rotational) magnetic field is generated by a time-dependent electric field and a current density **j**:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}, \tag{2.34}$$

where

$$\mu_0 = \frac{1}{(\varepsilon_0 c^2)} = 4\pi \times 10^{-7}\ \mathrm{H\ m^{-1}} \tag{2.35}$$

is the permeability of free space and the current density is defined as the current I per unit area.

2.7 Electromagnetic Energy Flow: Poynting Flux

The electromagnetic field carries energy. We need an appropriate equation that describes the energy balance in an EM-field. Starting with the Faraday Law described by Eq. (2.33) we can obtain the so-called electromagnetic field energy equation. First multiply the left-hand side of the Faraday's Law with the magnetic field and integrate over volume

$$-\int_V \mathbf{B}\cdot(\nabla\times\mathbf{E})\,\mathrm{d}^3r=\int_V \mathbf{B}\frac{\partial\mathbf{B}}{\partial t}\,\mathrm{d}^3r, \tag{2.36}$$

where the divergence of a curl is given by the identity

$$\nabla\cdot(\mathbf{a}\times\mathbf{b})=\mathbf{b}\cdot(\nabla\times\mathbf{a})-\mathbf{a}\cdot(\nabla\times\mathbf{b}). \tag{2.37}$$

So, the divergence of the vector product of the electric field with the magnetic field can be written as

$$\nabla\cdot(\mathbf{E}\times\mathbf{B})=\underbrace{\mathbf{B}\cdot(\nabla\times\mathbf{E})}_{=1}-\underbrace{\mathbf{E}\cdot(\nabla\times\mathbf{B})}_{=2}. \tag{2.38}$$

If Ampere's Law (2.34) for the expression 2 is used, then Eq. (2.36) becomes

$$\mathbf{B}\cdot(\nabla\times\mathbf{E})=\nabla\cdot(\mathbf{E}\times\mathbf{B})+\mu_0\,\mathbf{E}\cdot\mathbf{j}+\mu_0\varepsilon_0\,\mathbf{E}\cdot\frac{\partial\mathbf{E}}{\partial t} \tag{2.39}$$

Using the divergence theorem

$$\int_V \nabla\cdot\mathbf{a}\,\mathrm{d}V=\oint_S \mathbf{a}\cdot\mathrm{d}\mathbf{S} \tag{2.40}$$

and the property

$$\mathbf{a}\cdot\frac{\partial\mathbf{a}}{\partial t}=\frac{1}{2}\frac{\partial(\mathbf{a}\cdot\mathbf{a})}{\partial t} \tag{2.41}$$

Eq. (2.36) becomes

$$\int_V\left(\mathbf{B}\frac{\partial\mathbf{B}}{\partial t}+\varepsilon_0\mu_0\,\mathbf{E}\cdot\frac{\partial\mathbf{E}}{\partial t}\right)\mathrm{d}^3r=-\int_S \mathbf{E}\times\mathbf{B}\cdot\mathrm{d}\mathbf{S}-\int\mu_0\,\mathbf{E}\cdot\mathbf{j}\,\mathrm{d}^3\mathbf{r} \tag{2.42}$$

$$\Rightarrow\frac{\partial}{\partial t}\int_V\left(\underbrace{\frac{B^2}{2\mu_0}}_{\varepsilon_B}+\underbrace{\varepsilon_0\frac{E^2}{2}}_{\varepsilon_E}\right)\mathrm{d}^3\mathbf{r}=-\int_S\underbrace{\frac{\mathbf{E}\times\mathbf{B}}{\mu_0}}_{\mathbf{S}_p}\cdot\mathrm{d}\mathbf{S}-\int_V\mathbf{E}\cdot\mathbf{j}\,\mathrm{d}^3\mathbf{r}, \tag{2.43}$$

where $\varepsilon_B = B^2/2\mu_0$ is the magnetic energy density, $\varepsilon_E = \varepsilon_0 E^2/2$ is the electric field energy density,

$$\mathbf{S}_p = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} \tag{2.44}$$

is the energy flux vector or the Poynting flux vector, and $\mathbf{E} \cdot \mathbf{j}$ is the Ohmic losses. $\mathbf{E} \cdot \mathbf{j} > 0$ represents loss and $\mathbf{E} \cdot \mathbf{j} < 0$ is a generator. Additionally, work is done by matter as $\mathbf{j} \times \mathbf{B}$, meaning that a field exerts a force on particles.

The Poynting flux can be interpreted as a measure of electromagnetic energy transfer to a system. For space science applications, this flux is formulated in terms of a convection electric field $\mathbf{E}_c$ and the perturbation of the geomagnetic field $\mathbf{B}'$ as $\mathbf{S}_p = \mathbf{E}_c \times \mathbf{B}'/\mu_0$. The solar UV/EUV portion of this energy flux impacts the upper atmosphere while passing through the magnetosphere and entering the thermosphere. The Earth-directed Poynting flux component can be retrieved from satellites such as the DMSP (Defense Meteorological Satellite Program) satellite (Knipp et al. 2011). The estimated values are up to 15 mW m^{-2}, being highly dependent on the Interplanetary Magnetic Field (IMF), and can reach 100 mW m^{-2} during magnetically disturbed conditions.

The most obvious manifestation of the electromagnetic energy is the light, which itself is an electromagnetic wave composed of a spectrum of frequencies, which I shall discuss next.

2.8 Electromagnetic Spectrum and Photons

Light is an electromagnetic wave. Its interpretation as an electromagnetic wave has not been established until Maxwell who proposed that the time-varying magnetic fields generate electric fields ($\partial\mathbf{B}/\partial t \rightarrow \mathbf{E}$) and time-varying electric fields are a source of magnetic fields ($\partial\mathbf{E}/\partial t \rightarrow \mathbf{B}$). The electric and magnetic fields sustain each other, forming the basis for an electromagnetic wave, for example, the visible light. An electromagnetic wave carries energy and momentum and do not require a medium, while, for example mechanical waves do. Though, electromagnetic waves and mechanical waves can be described using similar mathematical techniques.

Figure 2.5 schematizes the electromagnetic spectrum, including the wavelength, frequency [Hz], and energy [eV] range, which covers a broad range of waves from the low-frequency radio waves ($\lambda \sim 10$ m) to the high-frequency γ-rays ($\lambda \sim 10^{-6}$ nm), demonstrating an amazing 15 orders of magnitude scale variation within the spectrum.

The visible spectrum, that is, the portion of the electromagnetic spectrum that can be directly seen by the human eye, extends from the familiar colors blue to red, corresponding to wavelengths from ~380 to 750 nm. The frequency and wavelength are inversely related as

$$c = \frac{\lambda}{T} = \lambda f. \tag{2.45}$$

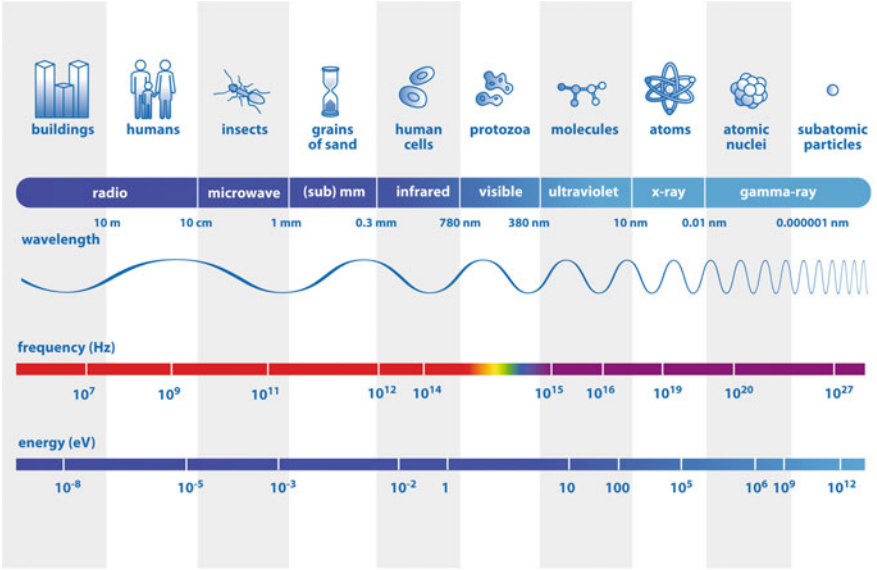

Fig. 2.5 A scheme of the electromagnetic spectrum with indication of wavelengths [nm], frequencies [Hz], and energies [eV] covering a spectrum of radio waves to gamma rays. The visible range illustrated with different color scales span wavelengths from 380 to 750 nm. Copyright: ESA / AOES Medialab

Thus the visible spectrum have frequencies ~7.9–4 $\times 10^{14}$ Hz. This implies that a photon, which is a part of the visible spectrum completes 10^{14} oscillations per second, which the human eye perceives as a blue to red color, although we are not capable of localizing individual photons. The energy of an electromagnetic wave with a given frequency (or wavelength) can be determined from

$$E = hf, \tag{2.46}$$

where $h = 6.626 \times 10^{-34}$ J s is the Planck constant. This relation is a consequence of Planck's postulate, with regard to the cavity radiation problem, that an oscillator cannot have continuous energies but only discrete energies specified by a number called the quantum number, n. His second postulate was that the oscillators do not radiate energy continuously, but only in certain steps or levels. When a change of state from one quantized level to another takes place, then energy is radiated that is given by the quantum number changes $\Delta E = \Delta nhf$, Following his postulates, Planck has derived a function that expresses the blackbody intensity as a function of temperature and frequency (wavelength) as

$$B_f = \frac{2hf^3}{c^2}\left[\exp\left(\frac{hf}{kT}\right) - 1\right]^{-1}. \tag{2.47}$$

Light as an electromagnetic wave is quantized. Einstein postulated that a beam of light consists of small packages of energy called *photons* or *quanta*. Equation (2.46) thus represents the energy of a single photon. One has to be cautious about the particle interpretation of light because photons are not particles in the usual sense. They travel in vacuum at the speed of light, have zero rest mass, possess momentum $p = E/c$, and demonstrate wave-like characteristics. The phenomenon of photons demonstrating both particle and wave characteristics is known as the *wave-particle duality*, which has caused an extensive amount of controversy in the development of quantum mechanics before the Second World War. However, the *complementarity principle* has been introduced, which was originally adopted from the German expression "Komplementaritätsprinzip", has allowed the application of both models without a contradiction.

According to (2.46), high-frequency waves, such as ultraviolet waves, have larger energy than low-frequency waves, such as the infrared waves. The energy emitted or radiated by electromagnetic waves is called the *electromagnetic radiation*.

So, how much is the energy of a single photon? In fact, the energy of a single photon is extremely small. For example, a photon belonging to the visible spectrum has an energy range of $\sim$2.65–5.24 $\times 10^{-19}$ J. Given that 1 eV is 1.602×10^{-19} J we get a range of 1.65–3.37 eV. The solar radiation is the most obvious example of electromagnetic radiation although maybe we do not appreciate it on a day-to-day basis, as we are used the natural day-night cycle. Sun is the source of natural light.

The solar spectral irradiance (SSI) or the solar spectrum is shown in Fig. 2.6 as observed by NASA's SORCE spacecraft launched in 2003. The primary goal of the SORCE mission is to precisely measure solar radiation. It provides spectral irradiance measurements in the range 1–2000 nm, which accounts for $\sim$95% of the spectral contribution to the total solar irradiance (TSI). The SSI plot presents the solar irradiance expressed in W m^{-2} nm^{-1} as a function of wavelength in nm. It is seen that the spectral energy density peaks around 500 nm, i.e., within the visible range of the spectrum, with about 2.1 W m^{-2} nm^{-1}. These values have also been reported in earlier texts (Ratcliffe 1972). Overall, the greatest amount of energy is in the range between 200 and 1200 nm. In this wavelength range, the average spectral energy density is $\sim$1.11 W m^{-2} nm^{-1} and the total energy flux amounts to $\sim$1110 W m^{-2}. Below 200 nm, the spectral density is very small (order of 10^{-5} W m^{-2} nm^{-1}), however, the short-wavelength part of the solar spectrum is the range that concerns the ionosphere the most, therefore its variability substantially impacts the ionosphere. The total energy flux of $\sim$1110 W m^{-2} calculated from the SORCE data is smaller than the *solar constant*

$$S_c = 1368 \text{ W m}^{-2}, \tag{2.48}$$

which is the average amount of radiative solar energy reaching the top of Earth's atmosphere at a mean distance of 1 AU ($= 149.6 \times 10^6$ km). The total energy output of Sun is by no means constant. It demonstrates variations over various time scales. The Total Irradiance Monitor (TIM) on board the SORCE spacecraft measures the radiant power across the entire solar spectrum from X-ray to far infrared scales. The

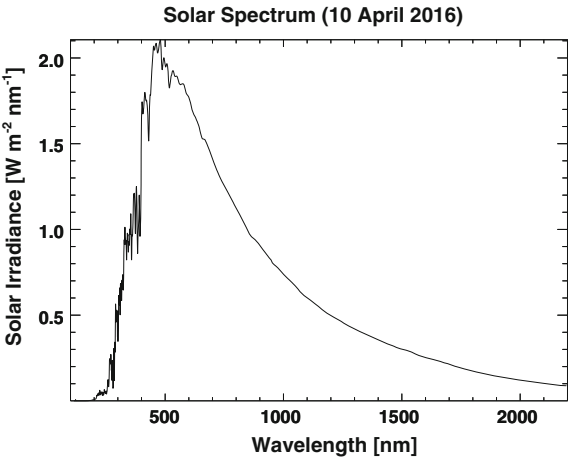

Fig. 2.6 Solar irradiance spectrum measured by the SORCE (Solar Radiation and Climate Experiment) spacecraft on 10 April 2016. The irradiance is plotted in W m^{-2} nm^{-1} as a function of wavelengths [nm]. The data is obtained from http://lasp.colorado.edu/lisird/sorce/sorce_ssi/

derived TSI was around 1361 W m^{-2} (Kopp et al. 2005), which was several W m^{-2} smaller than previous measurements.

The portions of the electromagnetic wave spectrum that are in particular relevant to the context of planetary atmospheres and ionospheres are the infrared (IR, $\lambda \sim 750$ nm–300 μm), ultraviolet ($\lambda \sim 10$–380 nm), and X-rays ($\lambda < 10$ nm) range, where the wavelength from 10–100 nm constitute the extreme ultraviolet portion (EUV). UV and EUV photodissociates O_2 in the terrestrial thermosphere. In particular, CO_2 and O_2 are strongly dissociated by solar radiation in the Schumann-Runge continuum (130–175 nm).

Figure 2.7 presents the penetration depth of solar radiation of wavelengths up to 300 nm in the terrestrial atmosphere as a function of altitude. While both UV and EUV contribute to energy absorption in the thermosphere, UV is the main contributor to heating in the middle atmosphere as it is primarily absorbed by ozone.

Figure 2.8 shows the Earth's atmospheric spectra in the visible and near-infrared range of the electromagnetic spectrum as observed by the OMEGA instrument (Table 1.2) onboard Mars Express spacecraft on 3 July 2003, as the spacecraft is on its long journey to Mars. The observed spectra is influenced in particular by the Pacific Ocean for this rotational configuration. It is seen that the spectra is dominated by the molecular species, such as molecular oxygen, water, carbon dioxide, ozone, and methane in the visible and infrared range.

2.9 Single Particle Motion in Electromagnetic Fields

When charged particles are in electric and/or magnetic fields they experience an external force as discussed earlier. We typically make some assumptions in order to simplify the charged particle motion in electromagnetic fields. The plasma density

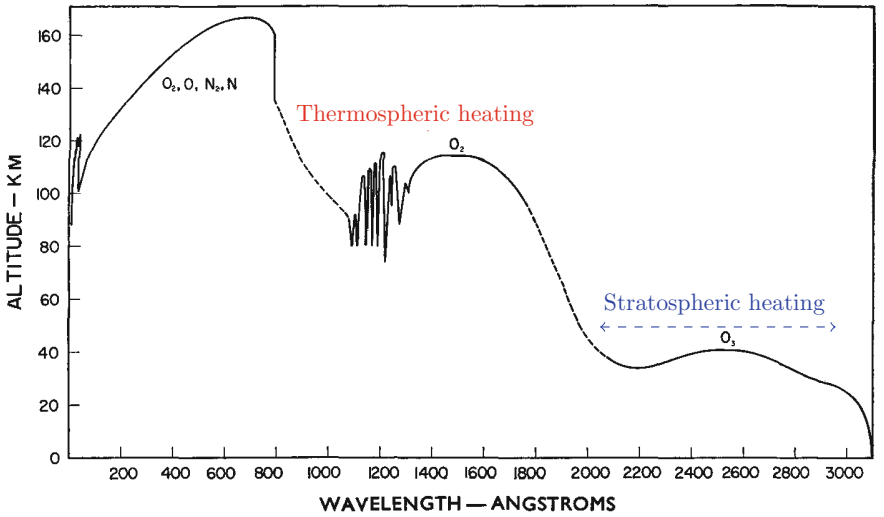

Fig. 2.7 Penetration depth of solar radiation in the terrestrial atmosphere. Photoabsorption by molecular oxygen above 100 km contributes toward thermospheric heating while photoabsoption by ozone between 30 and 50 km is the main source of stratospheric heating. Adopted from Friedman (1960)

is assumed to be small. Collisions between particles are ignored. The motion of the particles are determined only by the external fields and the energy density of the particles are small so that the external fields are not modified by charged particles.

One also has to consider the nature of the fields. Essentially, what one refers to as a field is a spatial distribution of a physical property. Typically, fields are two- or three-dimensional. For example, in a two-dimensional temperature field, the value may vary in two dimensions in space described by two coordinates, such as $T = T(\mathbf{r}) = T(x, y)$. Here temperature is a scalar field. The electric field $\mathbf{E} = \mathbf{E}(\mathbf{r})$ is a vector field. Fields can be time-dependent. In planetary atmospheres and ionosphere, we deal with time-dependent three-dimensional fields, e.g., $T = T(\mathbf{r}, t), \mathbf{v} = \mathbf{v}(\mathbf{r}, t)$, etc. Thus, the most sophisticated flavor of atmospheric models are three-dimensional time-dependent nonlinear general circulation models (GCMs). Such models calculate the radiative, chemical, and dynamical properties of planetary atmospheres. From the perspective of variability, one can overall distinguish between smooth and turbulent fields. Smooth fields vary weakly on the spatiotemporal scales of gyration, while stochastic processes have to be taken into account in a turbulent field because the associated variations have time scales comparable to or less than the time scales of gyration. Particle ensembles have to be considered instead of individual particles and transport equations are solved instead of the equation of motion. For example, transport processes in the interplanetary medium are typically turbulent.

One simple case of particle motion is in the case of uniform B field and zero electric field. If the magnetic field is uniform in space in the absence of an electric field $\mathbf{E} = 0$ then a given charged particle will experience a force due to the magnetic force only. The equation of motion is given by

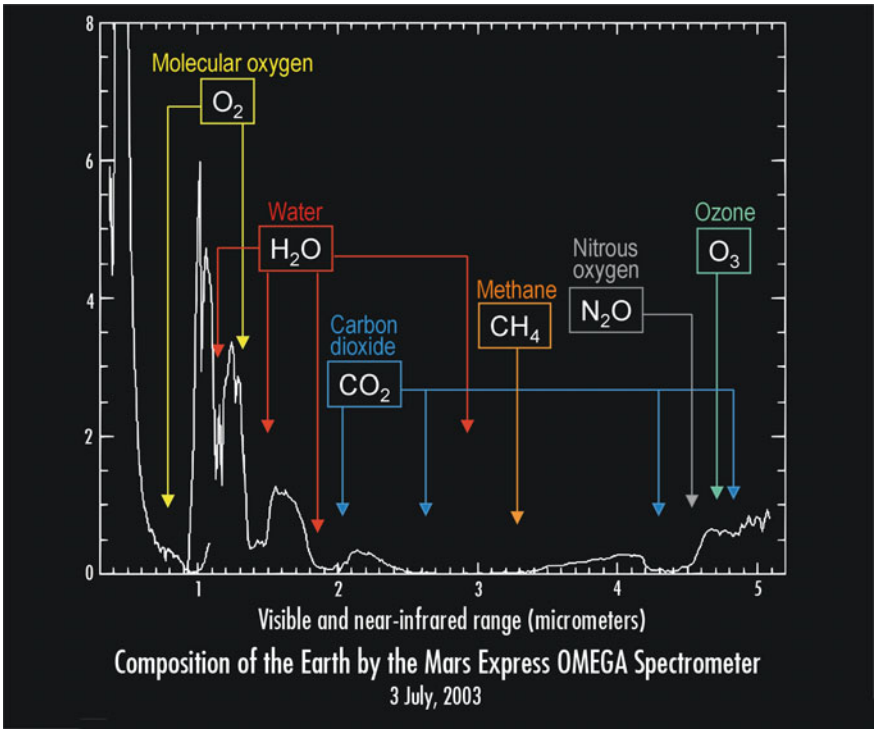

Fig. 2.8 Composition of Earth's atmosphere as detected by the Mars Express spacecraft in the visible and near-infrared range on 3 July 2003 on the way to Mars. The measurement was conducted by the OMEGA instrument facing Earth. The observed spectra is dominated by the Pacific Ocean and indicated the presence of molecular species, such as molecular oxygen, carbon dioxide, water, and ozone, and some other trace species like methane Credits: ESA/Institut d'Astrophysique Spatiale (Orsay, France)

$$\begin{aligned} \mathbf{F}_L &= \mathbf{F}_B \\ m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} &= q\mathbf{v}\times\mathbf{B} \end{aligned} \tag{2.49}$$

Let us assume that the magnetic field is directed along the z-axis as illustrated in Fig. 2.9, therefore $\mathbf{B} = (0, 0, B_z) = (0, 0, B)$ and $|\mathbf{B}| = B$. Then, the magnetic force is perpendicular to $\mathbf{B}$ and the components of the equation of motion are:

$$m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = q(v_y B, -v_x B, 0) \tag{2.50}$$

$$m\dot{v}_x = qv_y B \tag{2.51}$$

$$m\dot{v}_y = -qv_x B \tag{2.52}$$

$$m\dot{v}_z = 0 \tag{2.53}$$

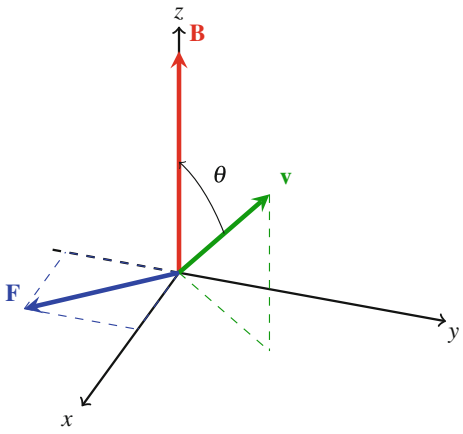

Fig. 2.9 Illustration of the magnetic field vector **B** which is aligned along the z-axis in this example, and a representative particle velocity vector **v**, which has components in all three coordinate directions, is shown in *green*. The angle between the magnetic field and the velocity is denoted by θ and the rotation of the velocity vector, $\mathbf{v} \times \mathbf{B}$, is shown by the *arrow*. The direction of the force vector **F** is marked in *blue* and is in the x-y-plane

The resulting force vector is on the xy-plane and is oriented in the positive x-direction and negative y-direction. This property can be deduced qualitatively from the components (2.51) and (2.52) of the equation of motion. Along the magnetic field no force acts on the particle and thus no acceleration occurs along the z-direction. Overall, the component of the velocity along the magnetic field, ($\mathbf{v}_{\parallel}$), leads to a drift along the magnetic field, while gyrating around it, owing to the perpendicular component of the velocity ($\mathbf{v}_{\perp}$), which experiences a magnetic force perpendicular to both the magnetic field and itself.

More complicated cases of magnetic field and electric field distribution are possible. For example, the electric field may be nonzero and the magnetic field may be nonuniform. In the presence of electric and magnetic fields, plasma drift manifests itself as a $\mathbf{E} \times \mathbf{B}$ drift. The exact nature of the electric field, whether it is time-dependent or uniform, will determine the $\mathbf{E} \times \mathbf{B}$ drift. This drift is given by

$$\mathbf{v}_{E\times B} = \frac{\mathbf{E} \times \mathbf{B}}{B^2}, \tag{2.54}$$

which is independent of the sign of the charge. Thus electrons and ions drift in the same direction, with electrons and ions being accelerated when moving antiparallel and parallel to the electric field, respectively.

2.10 Concept of Phase Space—Collection of Particles

The single particle motion treatment (Sect. 2.9) can be expanded to multiple particles by introducing the concept of *phase space* in which a collection of particles can be described by an associated distribution function that provides detailed information on the position and velocity of a given particle at time t. This description is more

appropriate because plasmas in nature consist of a collection of particles. In particular, large numbers of plasma particles have a spectrum of velocities.

The phase space consists of the configuration space **r** and the velocity space **v**. It is a six-dimensional space in time. That is, every particle (atom or molecule) is described by six coordinates in space: three coordinates denoted by x, y, and z, which are the positions of its center of gravity, and additional three coordinates of velocity components v_x, v_y, and v_z. The phase space is expressed by the generic distibution function

$$f(\mathbf{r}, \mathbf{v}, t), \tag{2.55}$$

which is the *phase space distribution function* for a single particle species. This distribution function can be used to describe the properties of plasma consisting of a large number of particles by stating how many particles are contained in the six-dimensional space (or volume element). The associated *phase space differential volume* $\mathrm{d}V_p$ is given by

$$\mathrm{d}V_p = \mathrm{d}|\mathbf{r}|\,\mathrm{d}|\mathbf{v}| = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}v_x\,\mathrm{d}v_y\,\mathrm{d}v_z \tag{2.56}$$

Multiplication of the phase space distribution function with the differential volume element represents the number of particles $\mathrm{d}N_p$ in the phase space differential volume

$$\mathrm{d}N_p = f(\mathbf{r}, \mathbf{v}, t)\,\mathrm{d}|\mathbf{r}|\,\mathrm{d}|\mathbf{v}|. \tag{2.57}$$

The total number of particles is then given by the integration of the phase space distribution function over all infinitesimally small phase space volume elements:

$$N_p = \int_{-\infty}^{+\infty} f(\mathbf{r}, \mathbf{v}, t)\,\mathrm{d}V_p. \tag{2.58}$$

2.11 Collisions

Collisions describe the process of two or more particles coming in contact with each other and exchanging various dynamical and thermodynamical properties. There are, in general, two basic simplified types of collisions in nature: *elastic* and *inelastic* collisions. Linear momentum $\mathbf{p} = m\mathbf{u}$ and total energy, kinetic plus potential $(E_k + E_p)$, are conserved for both, but for inelastic collisions kinetic energy is transformed to other forms of energy, e.g., heat, ionizations energy, etc., while for elastic collisions dissipative loss processes are negligible. In the discussion of collision, I will not go into the details of collision integrals, as these are beyond the scope of this book. Some rigorous treatment of collision integrals can be found, for example, in the books by Chapman and Cowling (1970) and Schunk and Nagy (2009). Thus, practical aspects of collisions will be discussed in our context. The force density (force per unit volume) experienced by the particles of species "i", $\mathbf{F}_{ij}$ due to collisions with

particles of species "j" is equal to the negative of the force density experienced by particles of species "j", $\mathbf{F}_{ji}$, due to collisions with particles of species "i"

$$\mathbf{F}_{ij} = -\mathbf{F}_{ji}. \tag{2.59}$$

This mutual force arises because of the relative velocity of the different particles with respect to each other. The force densities can then be conveniently expressed as

$$\mathbf{F}_{ij} = n_i\, m_i\, \nu_{ij}(\mathbf{u}_j - \mathbf{u}_i) = -n_i\, m_i\, \nu_{ij}(\mathbf{u}_i - \mathbf{u}_j), \tag{2.60}$$

$$\mathbf{F}_{ji} = n_j\, m_j\, \nu_{ji}(\mathbf{u}_i - \mathbf{u}_j) = -n_j\, m_j\, \nu_{ji}(\mathbf{u}_j - \mathbf{u}_i), \tag{2.61}$$

where n_i and n_j are the number densities; m_i and m_j are the masses of the particles; $\mathbf{u}_i$ and $\mathbf{u}_j$ are the drift velocities of the particle species; and ν_{ij} is the collision frequency for the collisions of the "i" particles with "j" particles, and ν_{ji} is the collision frequency for the collisions of the "j" particles with "i" particles. Often the collision frequency is interpreted as an effective momentum transfer coefficient and it is important to note that it is in general not a symmetric quantity, i.e., the frequency of the collisions of the "i" particles with "j" particles is not necessarily equal to that of the "j" particles with "i" particles, i.e., $\nu_{ij} \neq \nu_{ji}$. However, ν_{ij} and ν_{ji} are related to each other as

$$\rho_i \nu_{ij} = \rho_j \nu_{ji}, \tag{2.62}$$

where mass densities of the species are given by

$$\rho_i = m_i n_i, \qquad \rho_j = m_j n_j. \tag{2.63}$$

Considering collisions among various particles as a source/sink term in the momentum balance of a given particle species, we can write the associated time rate of change of velocity, resulting from collisional interactions. The acceleration of the "i" particles resulting from their collisions with the "j" particles is then

$$\left(\frac{\partial \mathbf{u}_i}{\partial t}\right)_{ij} = -\nu_{ij}(\mathbf{u}_i - \mathbf{u}_j). \tag{2.64}$$

Depending on the sign of the differential velocity, the collision force can lead to an acceleration or deceleration. We can solve the relation (2.62) for one of the collision frequencies

$$\nu_{ij} = \frac{\rho_j}{\rho_i}\nu_{ji}. \tag{2.65}$$

Therefore, in the case of the collisions of two particle species, the associated collision frequencies are equal to each other provided that $\rho_i = \rho_j$. In general, in planetary atmospheres there are multiple particle species, such as electrons, ions, and neutrals. Besides, there can be further species of neutrals and ions. Thus, a given particle species, let us say a specific ion species can have collision with all other existing ion

species, neutral compositional species, and electrons. One simplification that can be made in planetary atmospheres and ionospheres is that the particles of a particular species move with the same speed (see fluid model, Sect. 3.1), which implies that the members of the same species do not experience any collisions with each other, so in summary

$$\mathbf{F}_{ij} = \begin{cases} |\mathbf{F}_{ij}| = 0 \text{ if } i = j \\ |\mathbf{F}_{ij}| > 0 \text{ if } i \neq j \end{cases} \tag{2.66}$$

Such a simplification has great computational advantages in modeling of planetary atmospheres.

2.12 A Useful Mathematical Note: Summation and Product Notations

There is a large number of notations in physics. The summation notation "$\sum$" is one of the most common notations that is used, for example, to conveniently represent calculations with matrices and tensor. Given a variable $a_r (r = 1, 2, 3)$ and $x^r (r = 1, 2, 3)$. For example, a homogeneous linear function of the variables can be written in the form

$$\sum_{m=1}^{3} a_m x^m = a_1 x^1 + a_2 x^2 + a_3 x^3, \tag{2.67}$$

which depends only on one index "m", which runs from 1 to 3. This expression describes a system of first order (or simple systems) and the terms a_m and x^m are called the elements (or components) of the system. A second order system is defined by two indices, "m" and "n":

$$\begin{aligned} \sum_{m,n=1}^{3} a_{mn}\, x^m x^n &= \sum_{m=1}^{3}\sum_{n=1}^{3} a_{mn}\, x^m x^n \\ &= \left(\sum_{m=1}^{3} a_{m1}\, x^m x^1\right) + \left(\sum_{m=1}^{3} a_{m2}\, x^m x^2\right) + \left(\sum_{m=1}^{3} a_{m3}\, x^m x^3\right) \end{aligned} \tag{2.68}$$

which is a set of homogeneous quadratic function. Often, summations can be conveniently represented using the *summation convention* as

$$a_m x^m = a_1 x^1 + a_2 x^2 + a_3 x^3, \tag{2.69}$$

the left hand side of which is more compact than explicitly using the summation notation in (2.67). In the summation convention, a repeated small index is summed from 1 to 3. The repeated index is often referred to as the dummy index. For example, using the summation convention, the full system of linear equations can be conveniently represented by $a_{mn}\, x^m x^n = b_{mn}$.

The rules for the product notation are similar to the summation notation. For example,

$$\prod_{i=1}^{I} a_i = a_1 \cdot a_2 \cdot \cdots \cdot a_I. \tag{2.70}$$

For example, $n!$ (n factorial) can be expressed by the product notation

$$\prod_{i}^{n} i = n!. \tag{2.71}$$

2.13 Concluding Remarks

In this chapter I have summarized some basic aspects of electromagnetic theory, focusing on the electromagnetic waves and forces, currents, and collisions between particles. The next chapter is going to introduce transport equations, which are the basis of many fundamental equations relevant to planetary atmospheres and ionospheres.

References

Chapman S, Cowling TG (1970) The mathematical theory of non-uniform gases, 3rd edn. Cambridge Mathematical Library, Cambridge University Press, Cambridge

Friedman H (1960) The sun's ionizing radiation. In: Ratcliffe JA (ed) Physics of the upper atmosphere. Academic press, pp 378–456

Knipp D, Eriksson S, Kilcommons L, Crowley G, Lei J, Hairston M, Drake K (2011) Extreme poynting flux in the dayside thermosphere: examples and statistics. Geophys Res Lett 38(16). doi:10.1029/2011GL048302

Kopp G, Lawrence G, Rottman G (2005) The total irradiance monitor (TIM): science results. Springer, New York, pp 129–139. doi:10.1007/0-387-37625-9_8

Lorrain P, Koutchmy O (1998) The sunspot as a self-excited dynamo. Astron Astrophys 339:610–614

Ratcliffe JA (1972) An introduction to the ionosphere and magnetosphere. Cambridge University Press, Cambridge

Schunk RW, Nagy AF (2009) Ionospheres: physics, plasma physics and chemistry. Atmospheric and space science series. Cambridge University Press, Cambridge

Chapter 3
Transport Processes in Plasma

Fluid, Kinetic, and MHD Approaches

All good things which exist are the fruits of originality.
John Stuart Mill (1806–1873)

Abstract The goal of this chapter is to introduce the basic principles of various equations governing transport processes in plasma. A velocity distribution function is a convenient representation of the spatio-temporal characteristics of plasma. Using such a distribution function, plasma behavior can be simplified by considering fluid or kinetic models. In the fluid model, particles perform a collective motion and is thus described by a single temperature. On the contrary, the kinetic model treats the individual species separately, assigning them a different velocity and temperature. Various transport equations describing fluids can be derived from the Boltzmann equation. In particular ion diffusion is an important transport process in the ionosphere that influences the distribution of plasma. Electric conductivity in the ionosphere influences the amount of charge particle flow (current flow) for a given condition of ambient field. It is an important factor for the energetics of the ionospheric plasma and its coupling to the neutrals. The approach of magnetohydrodynamics (MHD) combines plasma physics with the fluid approximation. While most fluid models for the neutrals prescribe electric and magnetic fields, the MHD technique solves the field equations and the equation of motion simultaneously, self-consistently calculating electric and magnetic fields resulting from the interaction of particles with the fields.

Keywords Boltzmann equation · Maxwellian distribution · Transport equations · Ion diffusion · Electric conductivity · Stress tensor · Navier-Stokes equation · Magnetohydrodynamics

E. Yiğit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-62006-0_3

3.1 Introduction

Transport processes are of central importance in engineering aspects and fundamental physics of fluids. The general notion of transport processes, sometimes also referred to as transfer processes, describe the transport of energy, momentum, and mass. Mass transport can be understood as the transport of a homogeneous gas or as the transport of different species contained in a gas volume. For example, consider the transport of gas by diffusion, which can be represented by Fick's law in term of particle flux, which is proportional to a diffusion coefficient and the directed against the concentration gradient.

Two basic physical models are typically used to characterize the dynamics of fluids: the *fluid model* and *kinetic model*. In order to illustrate the difference between these models, let us revisit the notion of *velocity distribution* described by a corresponding *velocity distribution function* f (Sect. 2.10). Essentially, as the name suggests, the purpose of such a function is to characterize the distribution of velocities in space, that is, the configuration space. Then, $f_s = f(\mathbf{r}, t)$ is a function in the three-dimensional space and is time-dependent, as a given distribution can change over time. Such a description is possible when it is assumed that the velocity distribution of each species is Maxwellian (Sect. 3.5) everywhere and can therefore be described by a single temperature value. This picture qualitatively defines the fluid model. For example, a given volume element of neutrals or ions in the atmosphere may contain multiple neutral or ion species, which would be assigned a representative neutral or ion velocity, respectively.

However, in a more complex environment, a volume of gas may contain multiple species, denoted by "s" hereafter, which can have a spectrum of velocities, that is, each species is described by its own velocity distribution function: $f_s = (\mathbf{r}, \mathbf{v}_s, t)$, which describes the number of particles of species s that at time t are located in a configuration-space volume element $\mathrm{d}^3 r = \mathrm{d}V$ about $\mathbf{r}$ and have velocities in a velocity-space volume element $\mathrm{d}^3 v_s$ about the species velocity $\mathbf{v}_s$. Figure 3.1 illustrates the configuration-space (left) and the velocity-space (right), where the position and velocity vectors are shown in red. Therefore, following our discussion from Sect. 2.10, $f_s = (\mathbf{r}, \mathbf{v}_s, t)$ can be interpreted as a probability density function in the six-dimensional $(\mathbf{r}, \mathbf{v_s})$ phase space. In the fluid model, the independent variables are $\mathbf{r}$, t while in the kinetic model they are $\mathbf{r}$, $\mathbf{v_s}$, and t. Neither of the models describe the individual particles but a distribution of particles are considered. Here, the associated configuration-space and velocity-space volume elements are given by

$$\mathrm{d}^3 r = \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z, \tag{3.1}$$

$$\mathrm{d}^3 v_s = \mathrm{d}v_x\, \mathrm{d}v_y\, \mathrm{d}v_z. \tag{3.2}$$

As the velocity distribution function describes the positions and the velocities of individual species in a fluid, any changes in the positions and velocities in a gas species tells us how a system evolves in time and space. Thus, an interesting key question in this context is how we can describe the changes in the distribution function

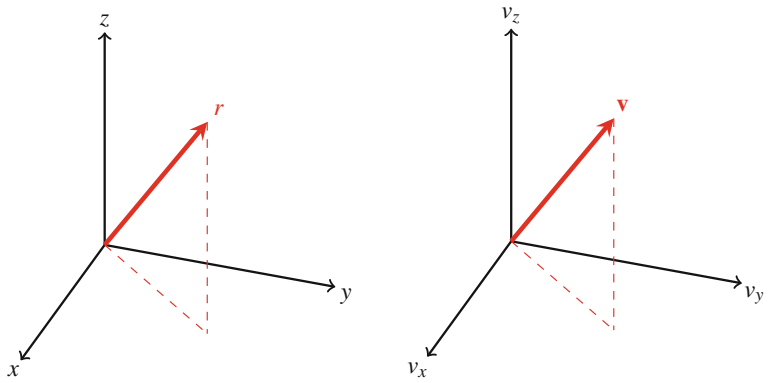

Fig. 3.1 Illustration of the vectors **r** in the configuration space (*left*) and **v** in the velocity space (*right*)

conveniently. For example, in the kinetic model a certain number of particles of a species s at time t in the configuration-space d^3r will have some velocities between $\mathbf{v}_s + \Delta\mathbf{v}_s$ and $\mathbf{v}_s$. So, how do the space- and velocity configuration look like at a later time $t + \Delta t$? Basically, how is the temporal evolution of f_s described? This question is addressed by the Boltzmann equation, which we discuss next.

Overall, this chapter is dedicated to the discussion of transport equations, starting with Boltzmann equation, and its various forms. Moments of distribution functions, Maxwellian distribution, ion diffusion, electric conductivities, stress, and Navier-Stokes equations will be discussed. A brief introduction to basic principles of magnetohydrodynamics (MHD) will be given.

3.2 Boltzmann Equation

We have essentially introduced the basic characteristics of the Boltmann's approach in the previous section. Namely, we are interested in the motions of a distribution of particles and not of individual particles. The gas mixture consists of different species and each species possess a separate velocity distribution function $f_s(\mathbf{r}, \mathbf{v}, t)$, which expresses the number of particles of a species s at time t located in d^3r and in d^3v_s. The evolution of the velocity distribution function is determined by the flow in phase space particles under the influence of external forces and by the net effect of collisions.

The time rate of change of f_s is given by

$$\frac{\mathrm{d}f_s}{\mathrm{d}t} = \lim_{\Delta t \to 0} \frac{f_s(\mathbf{r} + \Delta\mathbf{r}, \mathbf{v}_s + \Delta\mathbf{v}_s, t + \Delta t) - f_s(\mathbf{r}, \mathbf{v}_s, t)}{\Delta t}, \tag{3.3}$$

where $f_s(\mathbf{r}+\Delta\mathbf{r}, \mathbf{v}_s+\Delta\mathbf{v}_s, t+\Delta t)$ can be expanded in a Taylor series as

$$f_s(\mathbf{r}+\Delta\mathbf{r}, \mathbf{v}_s+\Delta\mathbf{v}_s, t+\Delta t) = f_s(\mathbf{r}, \mathbf{v}_s, t) + \nabla f_s \cdot \Delta\mathbf{r} + \frac{\partial f_s(\mathbf{r}, \mathbf{v}_s, t)}{\partial t}\Delta t + \nabla_{\mathbf{v}} f_s \cdot \Delta\mathbf{v}_s + O(\Delta\mathbf{r}^2, \Delta\mathbf{v}_s^2, \Delta t^2), \tag{3.4}$$

where ∇ and $\nabla_{\mathbf{v}}$ are the gradient operators in the configuration and velocity space, respectively, and the last term in Eq. (3.4) represent the order of terms that have been neglected in the expansion. Taking the limit $\Delta t \to 0$ in Eq. (3.3) along with the above Taylor expansion, the time rate of change of f_s becomes

$$\begin{aligned}\frac{\mathrm{d}f_s}{\mathrm{d}t} &= \frac{\partial f_s}{\partial t} + \nabla f_s \cdot \frac{\Delta\mathbf{r}}{\Delta t} + \nabla_v f_s \cdot \frac{\Delta\mathbf{v_s}}{\Delta t} \\ &= \frac{\partial f_s}{\partial t} + \mathbf{v}_s \cdot \nabla f_s + \mathbf{a}_s \cdot \nabla_v f_s,\end{aligned} \tag{3.5}$$

where $\mathbf{a}_s$ is the species acceleration vector. Note that the higher-order terms drop out as $\Delta t \to 0$. Two simplified interpretations of Eq. (3.5) are possible. If collisions are not important then the time rate of change of the velocity distribution function is zero: $\frac{\mathrm{d}f_s}{\mathrm{d}t} = 0$, yielding the *Vlasov equation*

$$\frac{\partial f_s}{\partial t} + \mathbf{v}_s \cdot \nabla f_s + \mathbf{a}_s \cdot \nabla_v f_s = 0. \tag{3.6}$$

If the collisions are important, then $\frac{\mathrm{d}f_s}{\mathrm{d}t} \neq 0$ and the species velocity undergoes instantaneous changes and thus particles appear in and disappear from the velocity-space volume element $\mathrm{d}^3 v_s$. In other words, collisions correspond to a production and loss terms for the species distribution function. Taking into account the effects of collisions yields the Boltzmann equation

$$\frac{\partial f_s}{\partial t} + \mathbf{v}_s \cdot \nabla f_s + \mathbf{a}_s \cdot \nabla_v f_s = \frac{\delta f_s}{\delta t}, \tag{3.7}$$

where $\delta f_s/\delta t$ denote the effects of collisions. Again, if collisions are neglected $\delta f_s/\delta t = 0$, Boltzmann equation gives the Vlasov equation (3.6).

Two main external forces in planetary atmospheres and ionospheres are the gravitational and Lorentz forces (Sect. 2.4) under which a species experiences the following the total acceleration $\mathbf{a}_s$

$$\mathbf{a}_s = \mathbf{a}_g + \mathbf{a}_L = \mathbf{g} + \frac{q_s}{m_s}(\mathbf{E} + \mathbf{v}_s \times \mathbf{B}), \tag{3.8}$$

where q_s is the species charge. It is important to note that in rotating planetary neutral atmospheres also fictitious forces arise, such as, the Coriolis and centripetal forces.

In most cases, the collision term $\frac{\delta f_s}{\delta t}$ can be evaluated analytically by solving the so-called Boltzmann collision integral, which describes particles scattered into a given region of velocity-space (production term) and particles scattered out of the same velocity-space volume element (loss term) (Schunk and Nagy 2009).

3.3 Moments of Distribution Functions

Ideally, one can solve the Boltzmann equation (3.7) for each species in a gas mixture in order to obtain the distribution function f_s. The shortcomings of this approach are that only for relatively simple situations can be treated and that it is restricted to obtaining information on a limited number of low-order *velocity moments* of the species distribution function. The practice of evaluating velocity moments is based on the procedure of multiplying the species distribution function f_s by powers or products of $\mathbf{v}_s$ and then integrating over all species velocities $\mathbf{v}_s$. The beauty of this approach is that various velocity moments then correspond to the well known field variables, such as, number density, temperature, drift velocity, etc. For example, if the distribution function is integrated over all velocities, it yields the number of particles in the configuration-space $\mathrm{d}^3 r$, which is technically the *species number density* $n_s(\mathbf{r}, t)$. We can express this more formally by

$$n_s(\mathbf{r}, t) = \int f_s(\mathbf{r}, \mathbf{v}_s, t)\, \mathrm{d}^3 v_s. \tag{3.9}$$

The number density is the zeroth velocity moment. The average (or drift) velocity of a species $\mathbf{u}_s(\mathbf{r}, t)$, $\langle \mathbf{v}_s(\mathbf{r}, t) \rangle$ is given by

$$\mathbf{u}_s(\mathbf{r}, t) = \frac{\int \mathbf{v}_s\, f_s(\mathbf{r}, \mathbf{v}_s, t)\, \mathrm{d}^3 v_s}{\int f_s(\mathbf{r}, \mathbf{v}_s, t)\, \mathrm{d}^3 v_s} = \frac{\int \mathbf{v}_s\, f_s(\mathbf{r}, \mathbf{v}_s, t)\, \mathrm{d}^3 v_s}{n_s(\mathbf{r}, t)}. \tag{3.10}$$

For highly collisional gas mixture, a single drift velocity $\mathbf{u}$ can be assumed because individual species drift velocities $\mathbf{u}_s$ and temperatures T_s do not significantly differ from the average drift velocity and temperature of the gas mixture. For example, in the lower atmosphere, neutral species are well-mixed and they all drift with the same speed. The above approach was one of the earlier methods of determining the transport properties of a given gas mixture (Chapman and Cowling 1970). However, in multi-component planetary atmospheres and ionospheres, it can be more appropriate to define the transport properties of a given species with respect to the average drift velocity of that species because large relative drift velocities can occur between interacting species. From (3.10) we can deduce that the average value of a field variable $\chi(\mathbf{v}_s)$, $\langle \chi_s(\mathbf{v}_s) \rangle$, is given by

$$\langle \chi_s(\mathbf{v}_s) \rangle = \frac{1}{n_s(\mathbf{r}, t)} \int \chi_s(\mathbf{v}_s)\, f_s(\mathbf{r}, \mathbf{v}_s, t)\, \mathrm{d}^3 v_s. \tag{3.11}$$

It is important to note that the definition of the higher-order velocity moments are not unique. For example, temperature is a measure of spread about some average velocity, which must be determined first. Therefore, higher-order velocity moments must be defined with respect to a reference average velocity. Usually calculation of an average value for a field variable depends on the choice of an interval, that is, over what scales averaging is performed. This concept is analogous to averaging performed in turbulence theory. For example, the temporal (or similarly spatial) average of the neutral wind velocity $\mathbf{u}$ is given with respect to some time interval between t and $t + T$, $\mathrm{d}t'$,

$$\overline{\mathbf{u}} = \frac{1}{T}\int_{t}^{t+T} \mathbf{u}(\mathbf{r}, t')\,\mathrm{d}t'. \tag{3.12}$$

For perturbations, associated with wave-like processes, it is instructive to perform the averaging over the period of oscillation as

$$\overline{u'^2} = \frac{1}{T}\int_{t}^{t+T} u'^2(\mathbf{r}, t')\,\mathrm{d}t', \tag{3.13}$$

where averaging over the square of the fluctuations ensure nonzero values, as $\overline{u'} = 0$. Given the species drift velocity, one can define a species random (thermal) velocity $\mathbf{c}_s$ as

$$\mathbf{c}_s = \mathbf{v}_s - \mathbf{u}_s. \tag{3.14}$$

Thermal velocity fluctuations can be interpreted as an instantaneous deviation of species velocities from an average (drift) velocity. Conceptually, this is identical to the definition of turbulent velocity fluctuations $u' = u - \bar{u}$.

As temperature is a measure of the average thermal velocity of species in a gas, we have the relation

$$\frac{3}{2}kT_s = \frac{1}{2}m_s\langle \mathbf{c}_s^2 \rangle, \tag{3.15}$$

and solving for species temperature yields

$$T_s = \frac{1}{3}\frac{m_s}{k}\langle \mathbf{c}_s^2 \rangle, \tag{3.16}$$

where the notation "$\langle\ \rangle$" consistently denotes an average as introduced in (3.11):

$$\langle \mathbf{c}_s^2 \rangle = \frac{\int \mathbf{c}_s^2\, f_s(\mathbf{r}, \mathbf{v}_s, t)\,\mathrm{d}^3 v_s}{n_s(\mathbf{r}, t)} = \frac{\int (\mathbf{v}_s - \mathbf{u}_s)^2\, f_s(\mathbf{r}, \mathbf{v}_s, t)\,\mathrm{d}^3 v_s}{n_s(\mathbf{r}, t)} \tag{3.17}$$

Some other useful velocity moments (for a species s) are temperature T_s, heat flow vector $\mathbf{q}_s$, stress tensor σ_s, and heat flow tensor Q_s (Schunk and Nagy 2009). The various physical parameters that make up the two common forms of approxi-

mations, the five-moment and 13-moment approximations to the transport equations are summarized in Table 3.1.

The partial pressure of a gas species depends on the species temperature:

$$p_s = n_s k T_s, \tag{3.18}$$

where n_s is the species number density as defined above and k is the Boltzmann constant. Eq. (3.18) is another form of the ideal gas law (Yiğit 2015, Eq. (2.10)):

$$pV = N_m R T, \tag{3.19}$$

where N_m is the number of moles and $R = 8.134\ \mathrm{J\ mol^{-1}\ K^{-1}}$ is the universal gas constant. Noting that the total mass $m = M N_A$ and $\rho = m/V$ we get a different form of the ideal gas law often used in planetary atmosphere models:

$$p = \rho \frac{R}{M} T, \tag{3.20}$$

where M is the molar mass. In a gas with total number of species S the resultant pressure is then

$$p = k \sum_{s=1}^{S} n_s T_s \tag{3.21}$$

For example, let us say there exist a gas mixture of CO_2 and O_2 in a given volume. Then, the resultant pressure in that mixture is produced by the contribution of all the motions of the molecules of both species. Formally, one can state

$$p = n_{\mathrm{co_2}} k T_{\mathrm{co_2}} + n_{\mathrm{o_2}} k T_{\mathrm{o_2}} = k(n_{\mathrm{co_2}} T_{\mathrm{co_2}} + n_{\mathrm{o_2}} T_{\mathrm{o_2}}). \tag{3.22}$$

In its most generic form, the concept of pressure (or the thermodynamic pressure) is the normal component of the total stress σ_s acting on the fluid of species s:

$$\sigma_s = -p_s \mathsf{I} + \tau_s, \tag{3.23}$$

stating that the total stress in fluids can be expressed in terms of the normal stress, which expresses the effect of the pressure, and the viscous (shear stress) τ_s, where I is the unit dyadic given by

$$\mathsf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{3.24}$$

such that $\mathsf{I} p_s = p_s \mathsf{I}$. The stress tensor is a measure of the extent to which the gas deviates from an isotropic character. A gas mixture, in which the different gas layers move in the same direction and the same speed is an example of a fluid of isotropic

Table 3.1 Plasma transport properties in the five-moment and 13-moment approximations. The number density n_s, three-dimensional drift velocity $\mathbf{u_s}$, and temperature constitute the five-moment approximation, the simplest approximation to transport equations

Transport variable		Five-moment	13-moment
Species number density	n_s	✓	✓
Species drift velocity	$\mathbf{u_s}$	✓	✓
Species temperature	T_s	✓	✓
Species heat flow vector	$\mathbf{q}_s$		✓
Species stress	σ_s		✓

motion. Then, no shear stress occurs between the layers and thus no interaction between the layers associated with shear stress. If, however, there is a difference between the individual gas layers in terms of speed and/or direction then, there exists a shear stress between the layers, and thus an interaction. If collisions are important in a gas mixture, then the diagonal terms in the stress tensor are important. It is important to note that various literature of fluid dynamics assumes different notation for the total stress, viscous stress, and the presure term. The stress tensor will be discussed in detail in Sect. 3.9.

3.4 Transport Equations

Some physically important velocity moments are the species number density n_s, drift velocity $\mathbf{u}_s$, temperature T_s, viscous stress tensor τ_s, and heat flow vector $\mathbf{q}_s$, as emphasized before. A transport equation is essentially a differential equation that determines the time evolution of a physically important field variable (referred to as a velocity moment in this context).

There are different forms of transport equations, depending on the variable under investigation. Tendencies (i.e., the local time rate of changes) of the relevant velocity moments are related to the known conservation laws. For example, the tendencies of the number density, drift velocity, and the pressure (or temperature) need to be evaluated for the conservation of mass, momentum, and energy, respectively. These conservation laws are also known as the equations of continuity, motion, and energy (Yiğit 2015). Essentially, these equations are the basis for the prognostic equations needed for atmospheric modeling. We can then appreciate the power of the Boltzmann equation (3.7) and of Boltzmann's approach because it provides a convenient way of deriving various transport equations (TE) that are essential to modeling of planetary atmospheres and ionospheres. The simplified receipe is as follows. We multiply both sides of the Boltzmann equation (BE) with an appropriate function of species velocity $\mathbf{v}_s$ and integrate over all velocities, meaning that the integration

variable is the three-dimensional velocity-space volume element $\mathrm{d}^3 v_s$ (3.2) as illustrated below

$$\int \mathrm{BE} \times F(\mathbf{v}_s)\,\mathrm{d}^3 v_s \rightarrow \mathrm{TE}. \tag{3.25}$$

Before embarking on the details of this formalism, it proves convenient to rewrite the Boltzmann equation (3.7) in a conservative form. This is achieved by using the mathematical property

$$\boldsymbol{\nabla} \cdot (f_s \mathbf{v}_s) = \mathbf{v}_s \cdot \boldsymbol{\nabla} f_s + f_s(\boldsymbol{\nabla} \cdot \mathbf{v}_s) = \mathbf{v}_s \cdot \boldsymbol{\nabla} f_s, \tag{3.26}$$

$$\boldsymbol{\nabla}_v \cdot (f_s \mathbf{a}_s) = \mathbf{a}_s \cdot \boldsymbol{\nabla}_v f_s + f_s(\boldsymbol{\nabla}_v \cdot \mathbf{a}_s) = \mathbf{a}_s \cdot \boldsymbol{\nabla}_v f_s. \tag{3.27}$$

With divergence-free flow and acceleration, Boltzman equation (3.7) becomes

$$\frac{\partial f_s}{\partial t} + \boldsymbol{\nabla} \cdot (f_s \mathbf{v}_s) + \boldsymbol{\nabla}_v \cdot (f_s \mathbf{a}_s) = \frac{\delta f_s}{\delta t}, \tag{3.28}$$

from which various transport equations can be derived using the receipe (3.25). For example, integrating the Boltzmann equation in the conservative form yields the evolution of the species number density, which is commonly known as the continuity equation:

$$\frac{\partial n_s}{\partial t} + \boldsymbol{\nabla} \cdot (n_s \mathbf{u}_s) = \frac{\delta n_s}{\delta t}. \tag{3.29}$$

To obtain the evolution of the species drift velocity we multiply the BE with $m_s \mathbf{c}_s$ and multiply over all velocities, which gives

$$\frac{\mathrm{d}_s \mathbf{u}_s}{\mathrm{d}t} = \frac{1}{\rho_s} \boldsymbol{\nabla} \cdot \sigma_s + \mathbf{g} + \frac{q_s}{m_s}(\mathbf{E} + \mathbf{u}_s \times \mathbf{B}) + \frac{\delta \mathbf{M}}{\delta t}, \tag{3.30}$$

where

$$\frac{\mathrm{d}_s}{\mathrm{d}t} \equiv \frac{\partial}{\partial t} + \mathbf{u}_s \cdot \boldsymbol{\nabla} \tag{3.31}$$

is the convective derivative for a given species, and

$$\frac{\delta \mathbf{M}}{\delta t} \tag{3.32}$$

is the collision term, and

$$\rho_s = n_s m_s \tag{3.33}$$

is the species mass density.

One shortcomings of the transport equations is that they do not constitute a closed set of equations. Appropriate distribution functions have to be assumed in order to

evaluate the collision terms. One important distribution function is the Maxwellian velocity distribution that can prove useful in the context of planetary atmospheres.

3.5 Maxwellian Distribution

In an ideal gas, a given molecule (particle) moving with the molecular speed v has the kinetic energy of $\frac{1}{2}mv^2$ and it is assumed to have no potential energy as if it were a free particle. The Maxwellian speed distribution, or sometimes also called the Maxwell-Boltzmann distribution, is a distribution function that describes systems in the equilibrium state. In this physical picture, various assumptions are made in order to simplify the molecular motion:

- The molecular size is much smaller than typical distance between molecules;
- Collisions rarely take place; particles perform random motion;
- Intermolecular forces are neglected.

This distribution expresses that the probability of encountering particles decreases exponentially as the energy increases.

The Maxwell distribution is given by

$$f_s^M(\mathbf{r}, \mathbf{v_s}, t) = A_s \exp\left[- \frac{\frac{1}{2}m_s(\mathbf{v_s} - \mathbf{u_s})^2}{kT_s} \right], \tag{3.34}$$

where the subscript "s" represents a given species, T_s is temperature, $\mathbf{v}_s$ is particle velocity, $\mathbf{u}_s$ is the species drift velocity, $k = 1.3807 \times 10^{-23}\,\mathrm{J\,K^{-1}}$ is the Boltzmann constant, m_s is the particle mass, and A_s is the amplitude term. A representation of the Maxwellian distribution for a random species is shown in Fig. 3.2. As the temperature increases, it is seen that the distribution broadens, shifting the peak probability to higher velocity values, but the magnitude of the peak probability decreases.

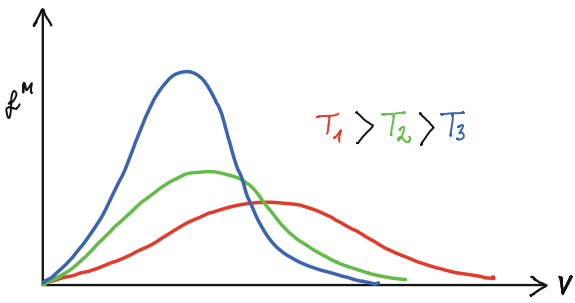

Fig. 3.2 Illustration of the Maxwell-Boltzmann distribution for three representative species temperatures: $T_1 > T_2 > T_3$. As the temperature increases the distribution broadens and the peak probability decreases

It is seen that the Maxwellian distribution includes the *Boltzmann factor* $\exp(-\varepsilon/kT)$, where

$$\varepsilon = \frac{1}{2}m_s c_s^2 = \frac{1}{2}m_s(\mathbf{v_s} - \mathbf{u_s})^2 \tag{3.35}$$

is the energy of the system. The specific shape of the curve is determined by the ratio ε/kT. The system of molecules are coupled to a larger system (or reservoir) at temperature T_s. The exponential characteristics of the Boltzmann factor suggests that for a given particle the probability of achieving a higher energy than kT_s is rapidly reduced. However, the particle has a reasonable chance of having an energy less than kT_s.

From the Maxwellian distributions one can determine what fraction of molecules are moving in a given velocity interval at a given temperature. The basic characteristics of a Maxwellian distribution are (1) the peak is given when the species velocity is equal to the species drift velocity; (2) the distribution is symmetric around the peak and it drops off exponentially away from the peak, and (3) the distribution decreases by a factor of "e" when $|\mathbf{c}_s| = |\mathbf{v}_s - \mathbf{u}_s| = (2kT_s/m_s)^{1/2}$. The width of the distribution is influenced by species temperature and the mass. The rms velocity is then given by

$$v_{rms} = \sqrt{\frac{3kT_s}{m_s}}. \tag{3.36}$$

3.6 Ion Diffusion

In terms of diffusive properties, ions in the ionosphere can be characterized in two groups: major and minor ions. We speak of a major ion when the ion density is comparable to the plasma density, i.e., $n_i \approx n_e$. A major ion thus have a substantial contribution toward maintaining charge neutrality in a plasma. A minor ion is identified as an ion species whose density is much smaller than the plasma density, i.e., $n_i \ll n_e$. It is a trace species and has negligible contribution to charge neutrality.

Diffusion of ions in the ionosphere is an important transport process as the F-region ionosphere, in particular the F_2 layer is not in photochemical equilibrium. Here I will briefly discuss two important ion diffusion processes: major ion (or ambipolar) diffusion and minor ion diffusion.

3.6.1 Major Ion Diffusion

In the case of major ion diffusion, it is assumed that the plasma consists of major ions, electrons, and only of one neutral species. The plasma is then partially ionized and heat flow, Coriolis and centripertal accelerations are neglected in the momentum

balance of ions and electrons. From the combination of the ion and electron momentum equations, including the above simplifications, neglecting the shear stress term, diffusion of ions along the magnetic field line, the so-called ambipolar diffusion equation, can be obtained:

$$\mathbf{u}_{i\parallel} = \mathbf{u}_{n\parallel} - D_a\left(\frac{1}{n_i}\nabla_\parallel n_i + \frac{1}{T_p}\nabla_\parallel T_p - \frac{1}{H_p}\right), \tag{3.37}$$

where the notation "$\parallel$" indicates the direction along the magnetic field; $\mathbf{u}_{i\parallel}$ then denotes the ion diffusive drift along the magnetic field; H_p is the plasma scale height

$$H_p = \frac{2kT_p}{m_i g}, \tag{3.38}$$

D_a is the ambipolar diffusion coefficient given by

$$D_a = \frac{2kT_p}{m_i \nu_{in}}, \tag{3.39}$$

with the ion-neutral collision frequency ν_{in} and ion mass m_i; T_p is the plasma temperature

$$T_p = \frac{T_e + T_i}{2}, \tag{3.40}$$

k is the Boltzmann constant.

Equation (3.37) has been the basis of previous theoretical planetary ionosphere models, for example in order to model the Martian ionosphere and interpret Viking 1 observations (Chen et al. 1978). Further simplifications to the ambipolar diffusion equation (3.37) is possible by assuming an isothermal plasma in thermal equilibrium with neutrals ($T = T_i = T_e = T_n$), which would eliminate the temperature gradient term.

The ambipolar diffusion equation applies to the configurations along the magnetic field in a strongly magnetized planetary ionosphere; also in the vertical direction for unmagnetized ionospheres. As the ion-neutral collision frequency is proportional to the neutral mass density, $\nu_{in} \propto n_n$ and $D_a \propto (\nu_{in})^{-1}$, the ambipolar diffusion coefficient decreases with increasing altitude. Neglecting the stress term and choosing r as the spatial coordinate either along the magnetic field in a strongly magnetized ionosphere or in the vertical direction in an unmagnetized ionosphere, one can obtain the classical diffusive equilibrium equation:

$$\frac{1}{n_i}\frac{\partial n_i}{\partial r} = -\frac{1}{H_p} - \frac{1}{T_p}\frac{\partial T_p}{\partial r}. \tag{3.41}$$

Integrating the classical diffusive equilibrium equation in altitude for an isothermal and constant gravity ionosphere yields

$$n_i(r) = n_i(r_0) \exp\left[\frac{-(r - r_0)}{H_p}\right]. \tag{3.42}$$

3.6.2 Minor Ion Diffusion

The minor ion diffusion can be obtained by making appropriate approximations to the momentum balance of ions. Inertial terms, the Coriolis force, the centripetal force and heat flow effects are neglected. The plasma is assumed to contain major ions, electrons, and neutrals, in addition to a given minor ion. The momentum exchange between the minor ions and the electrons is assumed to be small due to the small electron mass. Also, the electron-ion collision frequency is much larger than the electron-neutral collision frequency, $\nu_{ei} \gg \nu_{en}$, because of the long-range Coulomb effects. Under these physical assumptions and circumstances, the major ions can impact the motion of minor ions in three ways:

1. As major ions diffuse along the magnetic field, they tend to drag the minor ions with them,
2. In the process when the minor ions diffuse in response to their density and temperature gradients, collisions with major ions impact their motions,
3. A slight charge separation occurs between the major ions and electrons, which leads to a polarization electric field.

3.7 Electric Conductivities and Currents

Some basic principles of electric current and electric conductivity have been touched upon in Sect. 2.5. Essentially conductivity in plasma fluid, such as the partially ionized upper atmosphere, determines how much current can flow for a given electric field. We have to realize that electric conductivity can be significantly influenced by collisions with other particles.

Charged particle flow within a conducting plasma fluid (i.e., a conductor) depends on the ambient electric field and the physical properties of the plasma. Whenever, charged particles flow through the conductor they will experience a certain degree of resistance. This electric resistivity is the ratio of a given electric field to the current density:

$$\rho_r = \frac{|\mathbf{E}|}{|\mathbf{j}|}. \tag{3.43}$$

Thus, the greater the resistivity of a conductor, the greater is the electric field strength required to cause a given electric current density, or the smaller is the current for a given field strength. Accordingly, the conductivity is the inverse of resistivity

$$\sigma = \frac{1}{\rho_r} \tag{3.44}$$

and we get

$$\sigma = \frac{|\mathbf{j}|}{|\mathbf{E}|}, \tag{3.45}$$

yielding the simplified form of Ohm's law (2.24).

Currents in the ionosphere play an important role for the energetics and dynamics of the atmosphere-ionosphere (Anderson et al. 1998; Knipp et al. 2011). As discussed, a given resistivity (or conductivity) influences how much current can flow for a given electromagnetic field. Therefore, it is crucial to better understand ionospheric currents and the factors influencing them.

3.7.1 Equation of Plasma Motion

Following our previous discussion concerning collisions (Sect. 2.11) and electromagnetic force (Sect. 2.4) we can state equation of motion for a charged particle of species s drifting with velocity $\mathbf{u_s}$ under the influence of Lorentz force, neglecting all other body forces or external forces, while encountering collisions with other species denoted by j, which could be neutrals in a partially ionized plasma or ions in a fully ionized plasma,

$$\frac{\mathrm{d}\mathbf{u}_s}{\mathrm{d}t} = \frac{q_s}{m_s}(\mathbf{E} + \mathbf{u}_s \times \mathbf{B}) - \sum_j \nu_{sj}(\mathbf{u}_s - \mathbf{u}_j). \tag{3.46}$$

Current flow can be considered in general in the presence of electric and magnetic fields as indicated above. For illustrative purposes it is instructive to consider current flow in an unmagnetized plasma ($B = 0$) first.

3.7.2 Unmagnetized Plasma

Let us considered an unmagnetized steady-state plasma in which electrons are drifting with velocity $\mathbf{u}_e$ and all other collision partners are stationary. The momentum balance is given by the electric force and due to collision with the stationary neutrals:

$$-\frac{e}{m_e}\mathbf{E} = \nu_e\ \mathbf{u}_e. \tag{3.47}$$

where

$$\nu_e = \sum_n \nu_{en}, \tag{3.48}$$

considering that the momentum loss of electrons to ions is much smaller than their momentum loss to neutrals because in general the density of neutrals are much larger than ions in the ionosphere. Of course, it is implicit that, in a fully ionized environment such as the magnetosphere one would consider ions as collision partners.

Electron motion with respect to ions yields a current density given by

$$\mathbf{j}_e = -en_e\ \mathbf{u}_e \tag{3.49}$$

Combining Eqs. (3.47) and (3.49) gives the plasma resistivity

$$\rho_r = \frac{m_e\nu_e}{e^2 n_e}, \tag{3.50}$$

and thus the plasma conductivity

$$\sigma_0 = \frac{e^2 n_e}{m_e\nu_e}, \tag{3.51}$$

Therefore, a form of the generalized Ohm's law can be stated for unmagnetized plasma with near-infinity plasma conductivity as

$$\mathbf{j} = \sigma_0\ \mathbf{E}. \tag{3.52}$$

3.7.3 Magnetized Plasma

Presence of magnetic fields can significantly modify the electric currents and conductivities in a planetary ionosphere. For example, the terrestrial ionosphere is strongly influenced by the geomagnetic field, that is, it contains magnetized plasma ($B > 0$). The magnetic fields that are generated by current flows are too small and do not significantly affect the large-scale dynamics of plasma. Hence, the intrinsic magnetic field is taken to be well-known. If the ambient magnetic field were small, then Maxwell's equations (Sect. 2.6) would have to be solved coupled to the transport equations in order to estimate the contribution of the self-consistently generated fields.

If there exists a magnetic field and plasma drifts has a component across the magnetic field, then this effect has to be taken into account in the total generation of current density as

$$\mathbf{j} = \sigma_0(\mathbf{E} + \mathbf{u_s} \times \mathbf{B}), \tag{3.53}$$

Table 3.2 Major Ionospheric conductivities and the orientation of current flow with respect to the electric and magnetic field

Conductivity	Symbol	Orientation
Direct (or longitudinal) conductivity	$\sigma_\parallel$	$\mathbf{E}_\parallel, \mathbf{B}_\parallel$
Pederson (or transverse) conductivity	σ_p	$\mathbf{E}_\parallel, \mathbf{B}_\perp$
Hall conductivity	σ_H	$\mathbf{E}_\perp, \mathbf{B}_\perp$

where $\mathbf{u_s}$ denotes the drift velocity of the plasma of species s. This is an adequate description of current density in the fully ionized magnetosphere and solar wind (Baumjohann and Treumann 1997). However, the situation can be more complicated in the atmosphere-ionosphere system.

Due to a complex distribution of electric fields, magnetic fields, and neutral velocities, the conductivity is in fact an unisotropic tensor in the ionosphere. However, various simplifications can be made. For example, Rishbeth and Garriott (1969) discussed layer conductivity.

The formalism of obtaining the components of the conductivity tensor is in principle similar to the procedure for the unmagnetized case, but the equation of motion includes the magnetic fields. In general, ionospheric current flow occurs due to three types of conductivities, as summarized in Table 3.2, which are defined with respect to the mutual orientation of the electric and magnetic fields:

- Direct (or longitudinal) current $\sigma_\parallel$, which is parallel to the electric field and magnetic field ($\mathbf{E}_\parallel, \mathbf{B}_\parallel$);
- Pederson (or transverse) current σ_p, which is parallel to the electric field and perpendicular to the magnetic field ($\mathbf{E}_\parallel, \mathbf{B}_\perp$); and
- Hall current σ_H, which flows perpendicular to both fields ($\mathbf{E}_\perp, \mathbf{B}_\perp$)

The parallel, Pederson, and Hall currents are given by

$$\sigma_\parallel = \frac{n_e e^2}{m_e \nu_e} \tag{3.54}$$

$$\sigma_p = \sum_i \sigma_i \frac{\nu_i^2}{\nu_i^2 + \omega_{c_i}^2} + \sigma_e \frac{\nu_e^2}{\nu_e^2 + \omega_{c_e}^2} \tag{3.55}$$

$$\sigma_H = -\sum_i \sigma_i \frac{\nu_i\, \omega_{c_i}}{\nu_i^2 + \omega_{c_i}^2} + \sigma_e \frac{\nu_e\, \omega_{c_e}}{\nu_e^2 + \omega_{c_e}^2}, \tag{3.56}$$

where

$$\nu_i = \sum_n \nu_{in}. \tag{3.57}$$

and ω_{c_e} and ω_{c_i} are the electron and ion cyclotron frequencies, respectively, given by

$$\omega_{c_e} = \frac{|q|B}{m_e}, \qquad \omega_{c_i} = \frac{|q|B}{m_i}. \tag{3.58}$$

It is seen that the conductivities are appropriate for magnetized plasmas, which is the case in the lower ionosphere where the plasma is highly magnetized and collisional. Note that the parallel conductivity is equivalent to the plasma conductivity σ_0 in an unmagnetized case discussed in the previous section. For a conducting plasma moving with the velocity $\mathbf{v}$ in an electric field $\mathbf{E}$ across a magnetic field $\mathbf{B}$, the generalized Ohm's law describes the resulting current density

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{3.59}$$

$\mathbf{v} \times \mathbf{B}$ term is a consequence of the Lorentz transformation to (2.24) and σ is the anisotropic conductivity tensor given in components by

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} = \begin{pmatrix} \sigma_p & \sigma_H & 0 \\ -\sigma_H & \sigma_p & 0 \\ 0 & 0 & \sigma_\parallel \end{pmatrix}. \tag{3.60}$$

In Earth's atmosphere Hall conductivity peaks around 100 km, while the Pederson conductivity peaks at around 130 km. Often, it is convenient to use a height integrated conductivity, which can be formulated as

$$\Sigma_p = \int_{z_0}^{z} \sigma_p(z')\,\mathrm{d}z', \qquad \Sigma_H = \int_{z_0}^{z} \sigma_H(z')\,\mathrm{d}z', \tag{3.61}$$

where Σ_p and Σ_H are the height-integrated Pederson and Hall conductivities, respectively. Using conductivities and electric field distribution, the energy deposition rate associated with Joule heating [W m^{-3} = J m^{-3} s] can be estimated as

$$\mathscr{E}_J = \mathbf{j} \cdot \mathbf{E}. \tag{3.62}$$

The height-integrated energy deposition rate due to Ohmic losses can be expressed using the height-integrated conductivities. For example, the associated heating due to Pederson currents is

$$\mathscr{E}_{\Sigma_J} = \Sigma_p E^2. \tag{3.63}$$

Figure 3.3 presents the height-integrated Pederson conductivity, electric field strength, and height-integrated energy deposition by Joule heating as inferred by Aikio and Selkälä (2009) from the Tromso EISCAT UHF radar (69.59°N, 19.22°E) pointing in the field-aligned direction. Their results demonstrate the significance of the temporal correlation of conductivities and electric fields for the heating rates. The evening maximum is largely related to the electric field peak, while the morning

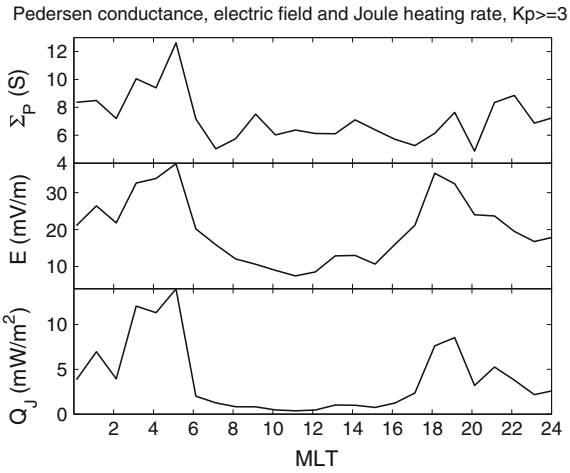

Fig. 3.3 Magnetic local time variations of the height-integrated Pederson conductivity Σ_p, electric field strength E, and energy deposition rate associated with Joule heating $\mathscr{E}_{\Sigma_J}$ (denoted as Q_J in the figure) as inferred by Aikio and Selkälä (2009) from EISCAT measurements in April 2006 for $K_p \geq 3$ conditions. The conductivity is integrated from 70 to 330 km

enhancement of Joule heating is shaped by both the conductivity and electric field peaks.

The total current density $\mathbf{j}$ is anisotropic and can be expressed in terms of components parallel and perpendicular to the magnetic field as

$$\mathbf{j} = \mathbf{j}_{\parallel} + \mathbf{j}_{\perp}. \tag{3.64}$$

In components, this is

$$\mathbf{j} = \sigma_{\parallel}\, \mathbf{E}_{\parallel} + \sigma_p\, \mathbf{E}_{\perp} + \sigma_H\, \hat{\mathbf{b}} \times \mathbf{E}_{\perp}, \tag{3.65}$$

or with $\hat{\mathbf{b}} = \mathbf{B}/B$ equivalently

$$\mathbf{j} = \sigma_{\parallel}\, \mathbf{E}_{\parallel} + \sigma_p\, \mathbf{E}_{\perp} - \sigma_H\, \frac{(\mathbf{E}_{\perp} \times \mathbf{B})}{B}. \tag{3.66}$$

3.8 Partial and Convective Derivatives

In hydrodynamics, one is in particular interested in changes in the dynamical field variables. For this, the use of the partial and convective (or advective) derivatives of a variables is a useful mathematical construct. The local (or partial) derivative $\partial/\partial t$ of a field variable expresses the time rate of change of that variable in a fixed location (thus local). The convective derivative $\mathbf{u} \cdot \boldsymbol{\nabla}$ describes the changes in a variable

because of the motion, represented by the drift velocity $\mathbf{u}$. If one is interested in a total change in a field variable, then we define a *total differential*

$$\frac{\mathrm{d}}{\mathrm{d}t} \equiv \frac{\partial}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla}, \tag{3.67}$$

which is a superposition of the local and the convective derivatives. In nature, all physical systems are dynamic. Nothing is static. Any change in a given field variable can be given by the superposition of the time rate of change in a fixed location and the changes due to a possible motion. In order to illustrate this property, imagine a weather station at an island. The local temperature changes can result from advective changes due to the motion of the air that can potentially transport either warmer or colder air from other regions around the island. Additionally, heating due to solar radiation can influence the local temperature. The total derivative is an essential part of the equation of motion. For example, the simplest representation of an equation of motion of a free particle i is the total derivative of the particle momentum, $\mathbf{p}_i$

$$\mathbf{F}_i = \frac{\mathrm{d}\mathbf{p}_i}{\mathrm{d}t} = \left[\frac{\partial}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla}\right]\mathbf{p}_i = \frac{\partial \mathbf{p}_i}{\partial t} + (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{p}_i. \tag{3.68}$$

The more familiar form is the force per unit mass (acceleration)

$$\frac{\mathbf{F}}{m} = \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u}. \tag{3.69}$$

The total derivative is used in the Lagrangian description of particle motion. For this, an infinitesimal volume element is assumed to move along the particle motion and various external forces are calculated on the moving volume element, following its motion.

Let us consider a volume element $\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ and assume that the properties of the fluid streaming through the volume element are measurable. Let the fluid property be represented by a function $f = f(x, y, z, t)$, which depends on the three Cartesian spatial coordinates x, y, and z and time t. This representation is an example of Euler's description of a fluid. Note that the spatial coordinates do not depend on time in the Eulerian approach. In the Langrange's description of a fluid, the variable describing the fluid properties is given by

$$f = f(\mathbf{r}, t) = f(x(t), y(t), z(t), t), \tag{3.70}$$

where the spatial coordinates depend explicitly on time. This representation is in accordance with the moving fluid element that moves with the flow. To illustrate the practical difference between the Eulerian and the Lagragian frameworks, consider the measurement of temperature with a thermometer. The temperature change that is measured by a stationary thermometer would be $\partial T/\partial t$, which would be according to the Eulerian system. However, if we consider a thermometer on a radio-sonde

carried by the prevailing wind, then the total time rate of change of temperature would be a superposition of changes in a fixed position and changes due to motion. In Euler's case, we have $\mathrm{d}f/\mathrm{d}t = \partial f/\partial t$. In Langrange's case, the total change in the fluid property can be expanded as

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t} = \frac{\partial f}{\partial t} + (\mathbf{u}\cdot\boldsymbol{\nabla})f, \tag{3.71}$$

which states mathematically that the total change in a moving fluid is given by changes at a fixed position (local temporal derivative) and by changes caused because of the relative motion between the observer and the medium (advective). Often, the total derivative is represented by D/Dt, which is identical to $\mathrm{d}/\mathrm{d}t$.

3.9 Stress Tensor

Forces that originate because of hydrodynamical stress is an important source of energy and momentum transfer in fluids. The total stress in fluids is composed of the thermodynamic pressure p and viscous forces, represented by the shear stress tensor τ (Sect. 3.3)

$$\sigma_{ij} = -p\delta_{ij} + \tau_{ij}, \tag{3.72}$$

where the total stress tensor σ_{ij} is a nine-component symmetric tensor in Cartesian coordinates (x, y, z). The diagonal terms of the stress tensor, σ_{xx}, σ_{yy}, and σ_{zz} include the pressure term and normal stress, while the off-diagonal terms represent the shear stress. For an inviscid fluid, the off-diagonal terms are zero.

The physical meaning of the indices are as follows: the first index represents the orientation of a given surface of a fluid parcel and the second index determines the direction of the stress. For example, σ_{zx} symbolizes stress acting in the x-direction across a xy-plane as illustrated in Fig. 3.4, or σ_{xy} is the stress terms in the y-direction across the yz-plane. The net force on a fluid parcel is given by the differences in the amount of stress acting across opposite surfaces. So, how do we determine the stress terms that contribute to a net force in a given direction? For example, if you take the x-direction, a net force would arise if there is a variation of the normal stress term, σ_{xx}, between the two opposing faces of the fluid parcel. Additionally, variations of the shear stress terms lead to a net force. Variations of σ_{yx} in the y-direction and of σ_{zx} in the z-direction would contribute to a net force in the x-direction. So, the force density (i.e., force per unit volume) in the x-direction is given by

$$\mathscr{F}_x = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{yx}}{\partial y} + \frac{\partial \sigma_{zx}}{\partial z} \tag{3.73}$$

Using the summation convention (2.69) one can write down the general expression for the total force density in the j-direction as

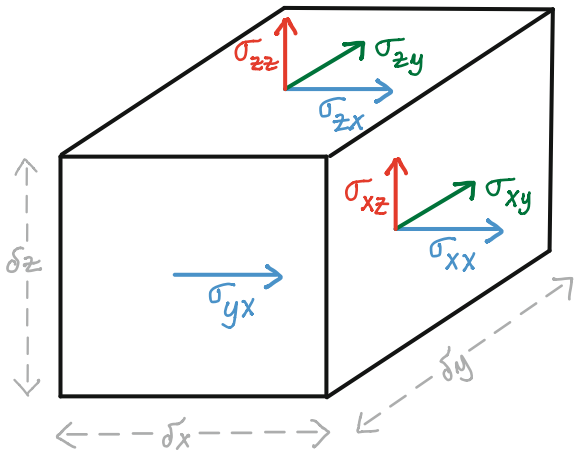

Fig. 3.4 Illustration of the components of the total stress tensor acting on a representative fluid parcel of volume $\delta V = \delta x \delta y \delta z$. The first index shows the orientation of the surface element, the second index shows the direction of stress. Different colors are used to indicate the stress terms acting in the same direction, such as, the σ_{xx}, σ_{yx}, and σ_{zx} terms, which would lead to a net force if there is a difference between the components acting across opposite faces

$$\mathscr{F} = \frac{\partial \sigma_{ij}}{\partial x_i} = -\frac{\partial p}{\partial x_j} + \frac{\partial \tau_{ij}}{\partial x_i}, \tag{3.74}$$

where

$$-\frac{\partial p}{\partial x_j} = -\frac{\partial p}{\partial x_j}\delta_{ij}, \tag{3.75}$$

with the Kronecker delta given by

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j. \end{cases} \tag{3.76}$$

A simplified momentum balance of a fluid parcel is achieved when gravity and the shear stresses are considered, which yields the Cauchy equation of motion.

3.10 Navier-Stokes Equations

The Navier-Stokes equation describes the momentum balance in a viscous fluid. Using the derivation of the total stress in Sect. 3.9 we can state a general form of the Navier-Stokes equation, including the Coriolis effect due to planetary rotation and gravity as

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = -\frac{1}{\rho}\nabla p + \frac{1}{\rho}\frac{\partial \tau_{ij}}{\partial x_i} - 2\boldsymbol{\Omega} \times \mathbf{u} + \mathbf{g}. \tag{3.77}$$

Depending on the approximation made to the viscosity terms, various forms of the Navier-Stokes equation may be assumed. In some studies, planetary atmospheres are assumed to be incompressible, hence $\nabla \cdot \mathbf{u} = 0$, the incompressible Navier-Stokes equation for rotating atmospheres is

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = -\frac{1}{\rho}\nabla p + \nu\nabla^2\mathbf{u} - 2\boldsymbol{\Omega} \times \mathbf{u} + \mathbf{g}, \tag{3.78}$$

where $\nu = \mu/\rho$ is kinematic viscosity, assumed to be constant in this simplification. Further sources of momentum deposition can be added to the right hand side of (3.78).

A practical approach to deal with the nonlinear Navier-Stokes equation is to split the flow $\mathbf{u} = (u, v, w)$ into a mean and a turbulent component, referred to as the *Reynolds decomposition*

$$\mathbf{u} = \bar{\mathbf{u}} + \mathbf{u}', \tag{3.79}$$

where the "overbar" denotes an appropriate averaging, for example, over time, so that the mean is constant over the averaging interval while the average of the fluctuation vanishes

$$\overline{\mathbf{u}'} = 0. \tag{3.80}$$

Accordingly for the individual fluctuating components of the flow we have

$$\overline{u'} = \overline{v'} = \overline{w'} = 0. \tag{3.81}$$

It is then imagined that turbulent motions act on the large-scale flow in a manner similar to the way microscopic (molecular) motions influence the macroscopic flow. In this sense, random (turbulent) velocity fluctuations lead to a net momentum transport within the fluid, which is conveniently represented by the Reynolds stress tensor

$$\psi_{ij} = -\rho_0 \overline{u_i' u_j'}, \tag{3.82}$$

where $\rho_0 = \bar{\rho}$ is the mean density and $\overline{u_i' u_j'}$ is the average of the correlation velocities, meaning that turbulent stress is acting in a mean turbulent flow in addition to the viscous and normal stress that the large-scale flow experiences. The Cartesian components of the Reynolds stress tensor are given by

$$\psi = -\rho_0 \begin{pmatrix} \overline{u'^2} & \overline{u'v'} & \overline{u'w'} \\ \overline{v'u'} & \overline{v'^2} & \overline{v'w'} \\ \overline{w'u'} & \overline{w'v'} & \overline{w'^2} \end{pmatrix}. \tag{3.83}$$

One can write ψ mathematically as a superposition of a symmetric ψ_S and anti-symmetric ψ_A components

$$\psi_{ij} = \frac{1}{2}(\psi_{ij} + \psi_{ji}) + \frac{1}{2}(\psi_{ij} - \psi_{ji}). \tag{3.84}$$

However, ψ_{ij} is symmetric, i.e., $\psi_{ij} = \psi_{ji}$, e.g., $\psi_{xz} = \psi_{zx}$, with off-diagonal components are the shear stress terms, while the diagonal terms represent the normal stress terms and we have

$$\overline{u'v'} = \overline{v'u'} \tag{3.85}$$

$$\overline{u'w'} = \overline{w'u'} \tag{3.86}$$

$$\overline{v'w'} = \overline{w'v'}. \tag{3.87}$$

If the fluctuations are isotropic, that is, they are the same in all directions, the off-diagonal components vanish and one speaks of an isotropic turbulence. The Reynolds tensor can be interpreted by considering momentum transfer produced by turbulent fluctuations. In general, ψ_{ij} can be understood as an average transport (or flux) of i-momentum in the j-direction. For example,

$$\psi_{xz} = \psi_{zx} = -\rho_0 \, \overline{u'w'}, \tag{3.88}$$

$$\psi_{vz} = \psi_{zv} = -\rho_0 \, \overline{v'w'} \tag{3.89}$$

represent the mean vertical flux of horizontal momentum in the atmosphere, typically used to parameterized momentum deposition by small-scale gravity waves. In this case, the divergence of the Reynolds stress $\partial \psi_{ij} / \partial x_i$ expresses momentum flux divergence, that is, transport of momentum into mean flow. The associated turbulent force per unit mass is, for example, in the j-direction is

$$a_j = \frac{1}{\rho_0} \frac{\partial \psi_{ij}}{\partial x_i}. \tag{3.90}$$

3.11 Introduction to Magnetohydrodynamics (MHD)

In studying the charged particle motion, one often prescribes some electric and magnetic fields and then evaluates the response of the particle motion to the known fields. It is assumed that the fields do not change their properties because of the particle

motion. In magnetohydrodynamics (MHD), we consider an ensemble (collection) of particles and electric and magnetic fields are not prescribed. They are determined by the positions and motions of particles. This formulation implies that the field equations and the equation of motion have to be solved simultaneously. For example, Maxwell equations (Sect. 2.6) would be field equations and the momentum equation is the equation of motion. Thus, in MHD one deals with the dynamical interactions between various field patterns with a set of particle trajectories. In other words, particles generate fields along their orbits, and field patterns force the particles to move on these orbits. This coupling between fields and particles constitute certainly a time-varying dynamical system. Despite the various complexities of the MHD theory, it can incorporate various simplifications to the physics of charged particles (plasma physics). For example, one can assume that all particles have the same speed, which is representative of the fluid theory. If a collection of particles are to be considered, then the fluid theory breaks down and the kinetic theory has to be assumed. As there can be a large number charged particle species, the kinetic theory could bear some complications. Imagine a plasma density of $n_e = 10^{12}$ m^{-3} in Earth's ionosphere. This means that there are 1 million particles in a cubic centimeter of volume, which illustrates how complicated the system can get. Thus, it is convenient to treat the plasma a fluid.

3.12 Basic MHD Equations

The Magnetohydrodynamics (MHD) is a powerful theory that is used to study the motion of plasma by treating it as a conventional fluid. A convenient starting point for the MHD is the one-fluid description of plasma. Let us first discuss some basic MHD assumptions.

Permeability and permittivity of free space, $\varepsilon = \mu = 0$: The medium can be neither magnetized nor polarized.

Low flow speeds, $u/c \ll 1$ **and** $c_p/c \ll 1$: Flow speeds are small compared with the speed of light c. An important physical consequence is that electromagnetic waves cannot be treated.

High conductivities, $E/B \ll 1$: Electrical conductivities are high, which cancels immediately strong electric fields. A consequence of this assumption is that the displacement current (as part of Maxwell equations) $\partial E/\partial t \sim 0$ compared to the induction current.

Linearity: The equations are linear in u/c, c_p/c, and E/B. Following this assumption, all higher order terms in these variables can be ignored.

Conservation laws: Conservation of mass, energy, momentum, and magnetic flux are required.

Then, the formal equations of basic MHD are given by the Maxwell equations (2.30)–(2.34), where in the Ampere's law the displacement current is ignored, and we have Ohm's law

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}), \tag{3.91}$$

the continuity of electric charge

$$\frac{\partial \rho_c}{\partial t} + \nabla \cdot (\rho_c \mathbf{u}) = 0, \tag{3.92}$$

the equation of momentum

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\,\mathbf{u} = -\frac{1}{\rho}\nabla p + \frac{\mathbf{j} \times \mathbf{B}}{\rho} + \mathbf{g} + \nu \nabla^2 \mathbf{u}, \tag{3.93}$$

and the equation of state

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{p}{\rho^\gamma}\right). \tag{3.94}$$

3.13 Summary of Transport Properties

Finally, I would like to briefly summarize the main characteristics of the approximations we have discussed in this chapter: the five-moment and the 13-moment approximation to the system of transport equations.

Five-moment approximation Stress and heat flow effects are neglected. The plasma properties are represented by five field variables the species number density n_s, the three-dimensional particle drift velocity $\mathbf{u}_s$, and the species temperature T_s. Anisotropic pressure distributions, thermal diffusion and conduction effects are not described. Stress and heat flow processes are neglected.

13-moment approximation In addition to n_s, $\mathbf{u}_s$, and T_s, the physically significant velocity moments are the stress tensor σ_s and the heat flow vector $\mathbf{q}_s$. The stress tensor includes the shear stress tensor and the thermodynamic pressure. Stress and heat flow processes are taken into account. Maxwellian velocity distribution is assumed. In the collision-dominated limit Navier-Stokes equations can be obtained.

Ionospheric currents are very important for the energy and momentum budget of the atmosphere-ionosphere system. The electric conductivity of the ionospheric plasma determines how much current can flow through the ionosphere for a given electric and magnetic field distribution. These conductivities vary as a function of time season, height, solar and geomagnetic activities and are strongly influenced by the neutral atmosphere.

Next chapter will focus on the properties of the terrestrial magnetic field and the structure of the ionosphere, focusing on Earth's and Mars' ionospheres.

References

Aikio A, Selkälä A (2009) Statistical properties of joule heating rate, electric field and conductances at high latitudes. Ann Geophys 27(7):2661–2673

Anderson BJ, Gary JB, Potemra TA, Frahm RA, Sharber JR, Wahington JD (1998) UARS observationos of Birkeland currents and Joule heating rates for the November 4, 1993, storm. J Geophys Res 103:26,323–26,335

Baumjohann W, Treumann RA (1997) Basic space plasma physics. World Scientific

Chapman S, Cowling TG (1970) The mathematical theory of non-uniform gases, 3rd edn. Cambridge Mathematical Library. Cambridge University Press, Cambridge

Chen RH, Cravens TE, Nagy AF (1978) The Martian ionosphere in light of the Viking observations. J Geophys Res 83(A8):647–664

Knipp D, Eriksson S, Kilcommons L, Crowley G, Lei J, Hairston M, Drake K (2011) Extreme poynting flux in the dayside thermosphere: examples and statistics. Geophys Res Lett 38(16). doi:10.1029/2011GL048302

Rishbeth H, Garriott OK (1969) Introduction to ionospheric physics. International geophysics series, vol 14. Academic press, New York

Schunk RW, Nagy AF (2009) Ionospheres: Physics, plasma physics and chemistry. Atmospheric and space science series. Cambridge University Press, Cambridge

Yiğit E (2015) Atmospheric and space sciences: neutral atmospheres, vol 1. Springer briefs in earth science. Springer, Netherlands. doi:10.1007/978-3-319-21581-5

Chapter 4
Planetary Ionospheres

Magnetic Fields, Chemical Processes, and Ionospheric Structure

I am only a physicist with nothing material to show for my labours. I have never even seen the ionosphere, although I have worked on the subject for thirty years. That does show how lucky people can be. If there had been no ionosphere I would not have been standing here this morning.

Sir Edward Appleton (1956)

Abstract An ionosphere is a consequence of photoionization and is the partially ionized portion of a planetary atmosphere. It contains free electrons and ions, whose dynamics produces complex current systems, depending on ambient electric fields, and collisional properties, and gyration around magnetic fields. In terms of the vertical extent, the ionosphere coincides with the thermosphere, where the neutral species are diffusively separated, owing to strong molecular diffusivity and insufficient turbulent mixing. In a similar fashion, the vertical profiles of plasma species are greatly influenced by diffusive equilibrium. Ionospheric dynamics is modified by the geomagnetic field, which has a strong dipole component. Chemical processes have to be considered in detail in additional to dynamics in the ionosphere. Therefore, in the chemical continuity equation, chemical production and loss processes cannot be neglected. A Chapman layer is a first order approximation of production of ionization by the absorption of solar energy, which is a good description for the lower ionosphere (E, F_1 regions), but higher up in the F_2 region transport becomes important. Qualitatively, Earth's and Mars ionospheres demonstrate overall various similarities, besides major differences. Interaction with other species, and external (solar effects) and internal (lower atmospheric) processes play an important role in both ionospheres.

Keywords Ionosphere · Diffusive equilibrium · Magnetic fields · Chapman layer · Space weather · Planetary ionosphere · Earth's ionosphere · Martian ionosphere

E. Yiğit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-62006-0_4

4.1 Introduction to the Ionosphere

The planets in the Solar System orbit around our Sun and they are subject to radiative effects. Therefore, planetary atmospheres are continuously affected by the solar radiation, leading to the formation of planetary ionospheres, which are typically partially ionized and interact with the background atmosphere in a complex manner. Further away from Sun, radiative and ionizing effects of Sun gradually decrease but the outer planets still receive substantial energy from Sun. In an ionosphere free electrons can exist over a short period of time and are captured by positively charged nearby ions because the electric force is attractive for unlike charges. Depending on ionizable constituents and the distance from Sun, ionospheres possess different characteristics. Two important aspects that planetary ionospheres differ from each other are the variations of the electron (or plasma) density and of the ionospheric species with altitude. The best known and observed ionosphere is the terrestrial ionosphere, which extends from ∼60 to 1000 km, depending on the solar and geomagnetic activities. During high solar activity the ionosphere is denser than during lower solar activity. This variation of the ionospheric depth is related to the fact that the morphology of the internal structure of the ionosphere is largely controlled by the UV and EUV portion of the solar radiation intensities, which demonstrate a distinct solar cycle variation (Rishbeth and Garriott 1969).

The existence of an ionosphere has also practical and technological implications. Many communication systems use the ionosphere as a layer to reflect radio signals. Thus, these communication signals can be transmitted over very long distances, which was first performed by Marconi by transmitting radio signals across the Atlantic between England and America in 1901. Appleton appreciated the property of the ionosphere to enable long-distance electromagnetic wave propagation with his paper "Wireless studies of the ionosphere" published in 1932 (Appleton 1932). Any disturbances in the ionosphere can impact the communication as well as navigation systems that use GPS (Global Positioning System) satellites around the world. During geomagnetic storms, when Sun is highly active and emits high energy solar wind impacting the magnetosphere, communication networks can experience severe disturbances.

In the ionosphere, chemical species are not well-mixed and are thus diffusively separated with increasing height. At greater heights, lighter species dominate (Sect. 4.2).

Earth possesses an intrinsic magnetic field. This magnetic field with a strong dipole component influences the motion of charged particles nearly at all heights in the upper atmosphere. In the *magnetosphere*, which is the region of the extension of Earth's magnetic field environment into space that is coupled to the ionosphere via electric fields, charged particles are predominantly influenced by the magnetic field (Sect. 4.3). The magnetic coordinate system is a convenient framework to describe the magnetic field and its orientation with respect to the geographical coordinate system (Sect. 4.4). Chemical processes and ionization influence the morphology of planetary ionospheres and their coupling to the neutral atmospheres (Sect. 4.5).

The Chapman layer approximation (Sects. 4.6–4.7) can help describe the vertical structure of planetary ionospheres, in particular, the regions of the ionosphere where transport processes play a minor role in comparison with photochemical processes, and thus ionization can be approximated by a simple solar control, such as the case in Earth's E and F_1 regions (Sect. 4.8) and in Mars' M1 and M2 layers (Sect. 4.9).

Space weather is broadly defined as collection of electrodynamical processes originating at Sun and influence Earth's surface, whole atmosphere system, and magnetic field environment. During magnetically disturbed conditions, space weather processes substantially impact the ionosphere and thermosphere via enhanced electric field convection and Joule and particle heating (Sect. 4.10).

Overall, this chapter describes the morphology, the physics and chemistry of planetary atmospheres, Earth's magnetic field, the ionosphere and basic ionospheric plasma processes, and Mars' ionosphere.

4.2 Diffusive Equilibrium

Earth's atmosphere is stably stratified, which means that a layer of lighter gas lies above a layer of denser gas. In fact, all known planetary atmospheres are stably stratified. When disturbances occur in the atmosphere, which may produce mixing between the neighboring layers, stable configuration is eventually reestablished despite the initial departure from equilibrium. A fundamental feature of all known stably stratified planetary atmospheres is that the temperature varies with altitudes. As discussed in Chap. 4 of the first volume of this book to some extent, various radiative and dynamical processes shape the vertical profile of the neutral temperature. The lower atmosphere, the troposphere, contains approximately 85–90% of the total atmospheric mass. This layer is the region where in general weather takes place. In the stratosphere and the mesosphere, which constitute the middle atmosphere, radiative and dynamical processes influence the temperature structure. Above in the thermosphere and ionosphere, solar radiation is the primary source of energy input; but lower atmospheric waves play an important role as well. In the lower and the middle atmosphere, the atmosphere with a composition of 78% nitrogen (N_2), $\sim$ 21% oxygen (O_2), 0.9% argon (Ar), and 0.03–0.04% carbon dioxide (CO_2) is overall turbulently mixed. That is, the relative portions of these neutral constituents are constant and thus the mean molecular mass m is constant

$$m = \frac{\sum_i n_i m_i}{\sum_i n_i} = \text{const.} \tag{4.1}$$

Despite the small CO_2 concentration (300–400 ppm) in the atmosphere, a small increase in its concentration has a great impact on the climate. Such an increased is typically explained by anthropogenic[1] activity.

[1] Caused by humans.

The pressure and density variations in the well-mixed lower and middle atmosphere are given by

$$p(z) = p_0 \exp\left(- \int_{z_0}^{z} \frac{\mathrm{d}z'}{H(z)} \right), \tag{4.2}$$

$$n(z) = n_0 \frac{T_0}{T} \exp\left(- \int_{z_0}^{z} \frac{\mathrm{d}z'}{H(z)} \right), \tag{4.3}$$

where p_0, n_0, and T_0 represent some reference values lower in the atmosphere, such as the surface values and $H = RT/Mg$ is a mean scale height. As the atmosphere is well-mixed the scale height and temperature are represented by single values for the atmosphere. From the above equations, the pressure and density variations can be directly related

$$p(z) = p_0 \frac{n}{n_0} \frac{T}{T_0}, \tag{4.4}$$

consistent with the ideal gas law.

Higher up in the atmosphere, around the mesosphere and lower thermosphere (MLT) a competition between perfect mixing and molecular diffusive mixing takes place. While turbulent mixing tries to keep the mean molecular mass constant, diffusion acts to establish conditions in which each chemical species is distributed according to its own pressure and temperature. At greater heights, in the thermosphere however, turbulent mixing becomes less important and molecular diffusion increasingly dominates. The hypothetical boundary between the turbulent region and the region that is dominated by molecular processes is referred to as the *turbopause* and is typically situated in the MLT at around ~105 km. Essentially, the turbopause is the height at which the turbulent (eddy) diffusion (κ_t) is comparable (or equal to) the molecular diffusion (κ_m). In a realistic atmosphere, the turbopause demonstrates seasonal and latitudinal variations (Offermann et al. 2007). The region below the turbopause is called the *homosphere* and the region above is the *heterosphere*.

Above the turbopause, in the heterosphere, each chemical species is distributed as though it were the only constituent. The vertical profile of each species depends on the specific scale height

$$H_s = \frac{R\, T_s}{M_s\, g}, \tag{4.5}$$

which depends on the temperature T_s and molar mass M_s of the species. For neutrals, the different species have the neutral temperature but the species molar masses can substantially differ. The primary reason for the occurrence of diffusive separation is due to a lack of turbulence at greater heights. Typically, turbulent processes become gradually less important than molecular processes above the *turbopause*. Therefore, for upper atmospheric processes, diffusive separation is very important.

Assuming that the height distribution of atmospheric gas obeys the ideal gas law

$$p = nkT, \tag{4.6}$$

where $n = m\rho$ is the number density and k is the Boltzmann constant, and the atmosphere is in hydrostatic balance

$$\frac{\mathrm{d}p}{\mathrm{d}z} = -\rho g, \tag{4.7}$$

then we obtain a definition of the *pressure scale height* H

$$-\frac{\mathrm{d}p}{\mathrm{d}z} = \frac{mg}{kT} = \frac{Mg}{RT} \equiv \frac{1}{H}, \tag{4.8}$$

where we have used the relations

$$M = mN_A \tag{4.9}$$

and

$$R = kN_A \tag{4.10}$$

The definition of the pressure scale height implies that lighter species have larger scale height. For example, the molar mass of oxygen molecule is $M_{\mathrm{O}_2} = 32\ \mathrm{g\ mol}^{-1}$, while the oxygen atom is two times lighter. The scale height of O_2 at around the turbopause ($\sim$105 km, $\sim 10^{-4}$ Pa, $\sim$180 K) would be about 4.8 km, while the atomic oxygen's scale height would be twice that value. So, at greater heights, this property implies that lighter species, i.e., species with smaller molar mass, increasingly dominate because of their relatively larger scale heights. These considerations are reflected by Eqs. (4.2) and (4.3), which we need to write for each species in the thermosphere, for example,

$$n_s(z) = n_{s_0} \frac{T_{s_0}}{T_s} \exp\left(-\int_{z_0}^{z} \frac{\mathrm{d}z'}{H_s(z)} \right). \tag{4.11}$$

4.3 The Terrestrial Magnetic Field

It has been known for a long time that Earth has some sort of an internal (i.e., intrinsic) magnetic field. This property means that Earth is able to generate and sustain a magnetic field of its own over large time scales. This magnetic field should be distinguished from the interplanetary magnetic field (IMF) that is an extension of Sun's magnetic field environment. In the outer regions of the magnetosphere at the magnetopause, IMF connects to Earth's magnetic field (see Fig. 4.3 of Yiğit (2015)). The first formal statement of Earth possessing a magnetic field can be found in Gilbert's work *De Magnete* in 1600, in which he explicitly described Earth as a giant magnet. Today, hundreds of observatories around the world monitor Earth's magnetic field.

What is the nature of magnetism? Can we see magnetism? Is it a macroscopic or microscopic phenomenon? Magnetism is caused by the interaction of moving electric charges. Magnetic forces influence only moving charged particles. In reality we have a coupling between moving charged particles and magnetic fields. Moving charges (i.e., current) produce **B**, and then **B** affects the moving charges as a result of a magnetic force $q\mathbf{v} \times \mathbf{B}$ (Sect. 2.3).

If a bar-magnet (permanent magnet) is free to rotate, it orients itself according to Earth's magnetic field: one end points north (the North Pole) and the other end points south (the South Pole). As in the case of electric charges, opposite magnetic poles (north-south or south-north) attract each other, and like poles (north-north or south-south) repel each other.

As indicated, Earth is a giant magnet. This does not mean that there is a magnet in the center of the planet. Actually, the geographic North Pole is close to the geomagnetic South Pole, which explains why the north pole of a compass needle shows geographic north. In space physics, the convention is that geomagnetic North Pole is actually the geomagnetic South Pole. As the geographic axis (i.e., axis of rotation) is not aligned with the magnetic axis, having a relative tilt of $\sim$11.5°, a compass needle does not point exactly north. It deviates slightly depending on the geographical location. This deviation is called the *magnetic declination* (D). The magnetic field is described conveniently in terms of the magnetic field lines, which are directed from the geomagnetic South Pole (i.e., actual geomagnetic North Pole), looping around Earth, and entering the geomagnetic North Pole (i.e., actual geomagnetic South Pole). It is important to note that the magnetic field lines are not the lines of force, i.e., they do not represent the direction of the magnetic force. We can appreciate this fact if we recall the introduction into magnetic force in Sect. 2.3. Namely, the magnetic force is perpendicular to both the magnetic field and the direction of charged particle motion. Magnetic field lines can have complex distribution and always form closed loops (no magnetic monopoles). The simplest scenario applies when they are uniform, which means that the magnitude and direction of the magnetic field is constant. Earth's magnetic field strength is in the order of 1 G (10^{-4} T).

Simplest model of Earth having a magnetic field can be achieved by imagining a giant rod magnet placed in the center of Earth. A more accurate representation of Earth's magnetic field would be based on an offset tilted dipole model, according to which the magnetic axis is offset by about 500 km from the center of Earth and it makes an angle of about 11.5 degrees with the geographic axis as illustrated in Fig. 4.1, in such a way that it goes through two points at Earth's surface. Given that Earth's radius is about 6380 km, the hypothetical bar magnet is relatively close to the center of the planet. Close to Earth's surface the dipole field approximation is pretty good. Then the associated dipole magnetic scalar potential is (Eleman 1973):

$$V_m = \frac{\mu_0}{4\pi} \frac{M_E \sin\theta_m}{r^3}, \tag{4.12}$$

where M_E is Earth's magnetic moment, θ_m is the magnetic latitude, and r is the radial distance. From the gradient of this potential the magnetic field can be derived as

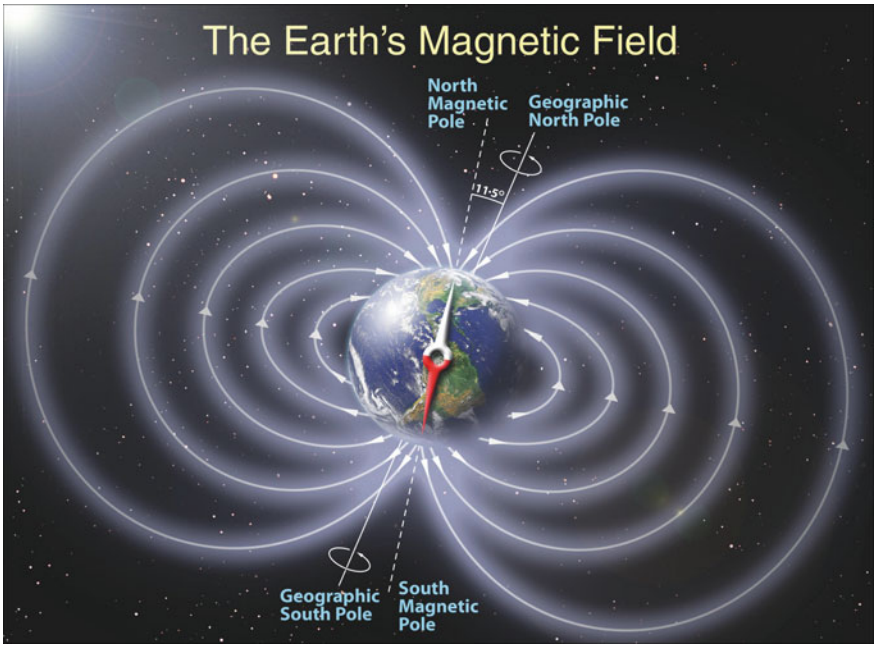

Fig. 4.1 Earth's magnetic field. Credit Peter Reid 2009

$$B(r, \theta, \phi) = -\nabla V(r, \theta, \phi), \tag{4.13}$$

with the components

$$\mathbf{B} = \frac{\mu_0 M_E}{4\pi r^3}(-2 \sin\theta_m \, \hat{\mathbf{r}} + \cos\theta_m \, \hat{\boldsymbol{\theta}}), \tag{4.14}$$

and the magnitude

$$B(r, \theta_m) = \frac{\mu_0 M_E}{4\pi r^3}\sqrt{(1 + 3 \sin^2\theta_m)}, \tag{4.15}$$

Away from Earth the dipole field approximation is not a good representation of the magnetic field in the magnetosphere. Earth's intrinsic magnetic field is generated in its core and possess appreciable spatio-temporal variations, which can be measured, for example, by ground-based magnetometers.

The measured field on Earth's surface is in general due to contributions from the main geomagnetic field, the Sq currents, field-aligned currents, ring currents, auroral electrojet, and the magnetopause currents. Typically, the main field varies much more slowly than the other fields (Gjerloev 2012).

At higher altitudes, measurements of Earth's magnetic field have been conducted with rocket-borne magnetometers in order to measure any departures from the inverse cube law (4.15) caused by currents generated in the ionosphere (Ratcliffe and Weekes 1960; Singer et al. 1951).

Earth's intrinsic magnetic field is adequately represented by empirical models. The International Geomagnetic Reference Field (IGRF) model that is currently in its 12-th generation ((IGRF-12), Thébault et al. 2015) is one of the widely used models. This model represents the spherical geometry of the time-dependent geomagnetic field $B(r, \theta, \phi, t)$ on and above Earth's surface produced by internal sources in terms of a scalar potential $V(r, \theta, \phi, t)$. In reality, the magnetic scalar potential is a superposition of an internal field that is due to core and lithospheric source and of an external field that describes mainly magnetospheric sources (Finlay et al. 2015):

$$V = V_{int} + V_{ext}. \tag{4.16}$$

The internal component can be readily expanded in terms of the so-called Gauss coefficients g_n^m and h_n^m and the Schmidt normalized associated Legendre Polynomial functions of degree n and order m, $P_n^m(\cos\theta)$ (Maus et al. 2005; Finlay et al. 2010).

$$V_{int}(r, \theta, \phi, t) = a \sum_{n=1}^{N} \sum_{m=0}^{n} \left(\frac{a}{r}\right)^{n+1} [g_n^m(t) \cos m\phi + h_n^m \sin m\phi] \times P_n^m(\cos\theta), \tag{4.17}$$

where r is the radial distance from the center of Earth and $a = 6371.2$ km is the so-called magnetic reference spherical radius, θ and ϕ are geocentric co-latitude and east longitude, respectively. In an analogous manner, the external field can be represented in terms of spherical harmonic expansion (Finlay et al. 2015).

4.4 Magnetic Coordinates

The geomagnetic coordinate system is defined with respect to the magnetic dipole axis. Various assumptions can be made concerning the orientation of this dipole axis with respect to the geographic axis in order to deal with Earth's intrinsic magnetic field

- A centered dipole field in which the geographic and the geomagnetic axes coincide;
- A tilted centered dipole field, making an angle of 11.5° with the geographical axis, with the dipole axis intersecting Earth's surface at two locations, 78.5° N 291° E and 78.5° S, 111° E; and
- An eccentric dipole, in which the tilted dipole field is offset by a distance of 500 km from the center of Earth.

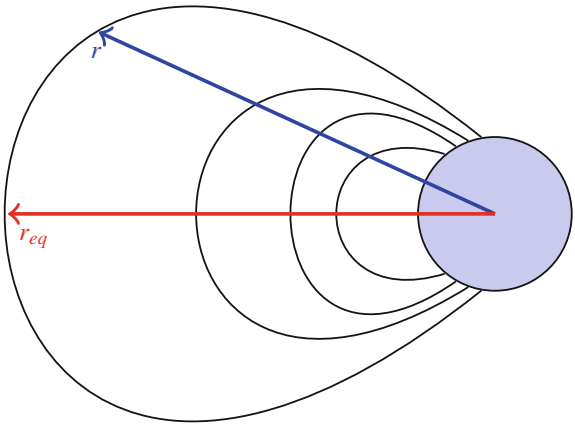

Fig. 4.2 Illustration of a dipole field around a planet, represented by a *blue circle*, where r marks the radial distance of a field line as a function of latitude and r_{eq} is the distance of the field line from the center of the planet above the equator

As discussed earlier, the most accurate representation of the magnetic field is obtained by expanding magnetic scalar potential in terms of spherical harmonic functions with certain expansion coefficients that can be determined by a network of ground-based and satellite borne magnetometer measurements. This approach is the basic principle of the International Geomagnetic Reference Field (IGRF) empirical model (Maus et al. 2005). More recently, Olsen et al. (2015) have derived a model of the Earth's magnetic field from the first year of Swarm satellite magnetic data.

The geomagnetic equator is defined by the intersection of the magnetic equatorial plane with Earth's surface. Overall the magnetic coordinate system is described by the geomagnetic longitude ϕ_m and latitude θ_m and the location of the boreal (Northern) magnetic pole are

$$\phi_{m_0} = 291^\circ \text{ E} \tag{4.18}$$

$$\theta_{m_0} = 78.3^\circ \text{ N}. \tag{4.19}$$

The magnetic and geographic coordinates can then be related by

$$\sin\theta_m = \sin\theta \sin\theta_{m_0} + \cos\theta \cos\theta_{m_0} \cos(\phi - \phi_{m_0}) \tag{4.20}$$

$$\sin\phi_m = \frac{\cos\theta \sin(\phi - \phi_{m_0})}{\cos\theta_m} \tag{4.21}$$

It is often convenient to describe the extension of Earth's magnetic field in terms of an L-shell parameter, which is a function of Earth's radius r_e

$$L = \frac{r_{eq}}{r_e}, \tag{4.22}$$

where r_{eq} is the distance of a given field line from the center of Earth as illustrated in Fig. 4.2. The surface corresponds to $L = 1$. The L-shell parameter is used for other Solar System planets by normalizing it for the planetary radius and planetary magnetic field model.

4.5 Chemical Processes

Planetary atmospheres include various chemical species, which can undergo various chemical and photochemical processes. These constituents can be present in atomic or molecular structures. Different planets have different chemical compositions and the chemical mixing is influenced to a different extent by dynamical processes. Earth has 75% N_2 while Mars' and Venus' atmospheres are composed of 95 and 96.5% CO_2, respectively. More exotic Solar System planets like Jupiter and Saturn are composed of more than 90% hydrogen. The ice giants, Uranus and Neptune, contain 83% and 80% H_2, respectively. Chemical processes occurring simultaneously with the dynamical processes greatly shape the spatiotemporal distribution of atmospheric neutral and plasma species.

Materials in the universe can be described in terms of their different kinds of *phases*. A phase in a material is the part which demonstrates the same properties and composition. For example, if a cube of ice is considered in a glass of full of water, then the glass contains two phases: water (liquid) and ice (solid). Chemical processes in the atmosphere can take place in gas phase or condensed phase or as a transition between these phases. If a chemical process occurs within the same phase then it is called a *homogeneous reaction*; if a phase transition occurs then it is a *heterogeneous chemical reaction*.

The major neutral constituent in the thermospheres of Venus and Mars is CO_2, while the dominant ionic species is O_2^+ because of ion-neutral chemical processes. On Earth, the atomic oxygen O is the major neutral constituent in the thermosphere and O^+ is the major ionic species. In order to provide some insight into the spectrum of reactions in planetary atmospheres let us first discuss some basic concepts of chemical processes.

4.5.1 Electronic Structure of Elements

An *element* is defined as a matter consisting of atoms that have nuclei with the same charge. All naturally occurring chemical elements are represented by the periodic table, in which each element is ordered according to the *atomic number* Z. This number represents the charge of the nucleus of an atom of an element, in units equal to the charge of the proton. For example for oxygen we have $Z = 8$ as it contains

Fig. 4.3 Frauenhofer spectrum as adopted by the German Post Service in stamp. The different Frauenhofer lines are denoted by letters A to G, which represent specific portions of the solar spectrum

eight protons. Its molecular mass is $m \sim 16$ u and its molar mass is $M \sim 16$ g mol^{-1}. The total charge of the protons in the nucleus is equal to $Z\,e$ and of the electrons is $-Z\,e$.

In daily life, we know that if a matter is heated up, it may radiate when sufficiently large temperatures are reached. For example, in the kitchen, when we turn on the electric heating plate, it may start to glow when it becomes very hot. When you turn on the lamp in the living room, you essentially are exploiting the high resistivity in the incandescent lamp, in which the wire or the filament within the lamp heats up and glows when charged particles flow through it. Essentially, it is a fundamental law of nature that all matter radiates depending on the temperature. In fact, there is a lot happening at atomic scales when matter receives energy. When energy is supplied to a system, individual atoms are excited. They may emit light, which can be diffracted to a specific pattern of discrete frequencies. This pattern is called a *line spectrum*. Frauenhofer was one of the first who studied such spectra. An example of a spectrum called the Frauenhofer spectrum, is seen in Fig. 4.3, which has been adopted by the German Post Service ("Deutsche Bundespost") in 1980s.

In the late 19th century scientists have started to study the line spectra of atoms. In 1913, Niels Bohr, a prominent quantum physicist, has provided the first successful interpretation of the hydrogen spectrum in terms of the electronic structure of the hydrogen atom. In contrary to the classical theory, which predicted a continuous spectrum for the hydrogen atom, experiments had demonstrated that the hydrogen spectrum consisted of discrete lines. Bohr postulated that the hydrogen atoms can exist only in discrete energy levels. That is, not all energies are possible. He defined the minimum energy level of a hydrogen atom as the *ground state*, which is the most stable state. The other energy levels have larger energies than the ground state and are called *excited states*. Then, light is emitted via photons if a transition occurs from an excited state (upper level) with energy E_U to a lower-level state with energy E_L

$$E_U - E_L = h\nu, \tag{4.23}$$

Table 4.1 Electron-dot symbol representation of the oxygen, carbon, and hydrogen elements

Oxygen ($^{16}_{8}$O)	Carbon ($^{12}_{6}$C)	Hydrogen ($^{1}_{1}$H)
$\cdot\dot{\underset{\cdot\cdot}{\mathrm{O}}}\cdot$	$\cdot\dot{\mathrm{C}}\cdot$	$\cdot\mathrm{H}$

which is referred to as the *Bohr frequency rule*. Equation (4.23) also applies in the absorption of light by atoms. The frequency of the light that is being absorbed in the transition from a lower level to an upper level is the energy difference between the upper and the lower level divided by the Planck constant

$$\nu = \frac{E_U - E_L}{h}. \tag{4.24}$$

Additionally, Bohr assumed that the electrons are on a circular orbit around the nucleus and that the associated angular momentum L is quantized:

$$L = mvr = n\hbar \qquad n = 1, 2, 3, ..., \tag{4.25}$$

where mv is the electron linear momentum, r is the orbital radius, n is the principal quantum number, and the

$$\hbar = h/2\pi \tag{4.26}$$

is the angular-momentum quantum. Bohr's atomic model has later been extended by the German physicist A. Sommerfeld who has included to some extent the elliptical orbits in the atomic structure (i.e., Bohr-Sommerfeld description).

The electronic structure of elements can be conveniently represented by an *electron-dot-symbol*. For example, the representation of the oxygen ($^{16}_{8}$O), carbon ($^{12}_{6}$C), and hydrogen ($^{1}_{1}$H) atoms are shown in Table 4.1. An electron in the outermost shell (valence electron) is represented by a dot. The distribution of electrons around a nucleus can be represented in terms of orbitals, s, p, d, etc. For example, the electronic configuration for the atomic oxygen is $1s^2\,2s^2\,2p^4$.

4.5.2 Chemical Kinetics

The field of chemical kinetics studies chemical reactions that cause the transformation of one or more kinds of matter into a new kind of matter. Such a chemical reaction can be represented by a general *stoichimetric equation*

$$\mathrm{aA} + \mathrm{bB} + \mathrm{cC} + \cdots \longrightarrow \mathrm{mM} + \mathrm{nN} + \mathrm{oO} + \cdots \tag{4.27}$$

where the species on the left hand side, i.e., A, B, etc., represent the *reactants*, and the species on the right hand side, i.e., M, N, etc., are the *products* of this chemical reaction, and the arrow "$\longrightarrow$" indicates the direction of the reaction, which is by convention from left to right for this one-way reaction. The coefficients, a, b, c, etc. describe how much of a given species is involved in the chemical reaction. A two-body reaction takes place between two species

$$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{C} + \mathrm{D} \tag{4.28}$$

For example, the *dissociative recombination* of an oxgen ion and and electron is an important loss process in the ionosphere:

$$\mathrm{O}_2^+ + \mathrm{e}^- \longrightarrow \mathrm{O} + \mathrm{O} \tag{4.29}$$

Chemical reactions can also be reversible, i.e., it can occur in both directions

$$\mathrm{O}^+ + \mathrm{H} \longleftrightarrow \mathrm{H}^+ + \mathrm{O}, \tag{4.30}$$

where *charge exchange* is taking place between O^+ and H or H^+ and O. So, the arrow "$\longleftrightarrow$" above indicates the reversibility of the reaction of the atomic oxygen ion and the hydrogen atom to the hydrogen ion and the atomic oxygen. When chemical products are formed directly from the reactants, one speaks of an elementary reaction.

If the recombination process is slow, other chemical reactions can occur during that process, which can be illustrated by considering the reaction of a generic atomic ion A^+ with a generic molecular species B_2 to an atom B and a molecular ion AB^+, which then undergoes a recombination process with an electron, yielding two atomic neutral species A and B:

$$\mathrm{A}^+ + \mathrm{B}_2 \longrightarrow \mathrm{AB}^+ + \mathrm{B} \tag{4.31}$$

$$\mathrm{AB}^+ + \mathrm{e}^- \longrightarrow \mathrm{A} + \mathrm{B} \tag{4.32}$$

For example, in the terrestrial ionosphere the recombination of the atomic oxygen ion with an electron is a slow process, thus it undergoes a *multi-step recombination process*, involving the nitrogen molecule, to ultimately form a nitrogen and an oxygen atom:

$$\mathrm{O}^+ + \mathrm{N}_2 \longrightarrow \mathrm{NO}^+ + \mathrm{N} \tag{4.33}$$

$$\mathrm{NO}^+ + \mathrm{e}^- \longrightarrow \mathrm{N} + \mathrm{O} \tag{4.34}$$

There are various ways to denote the concentration (or molar density) of a chemical species, which tells us how much of a given substance is contained in a unit volume. Typically,

$$[A] = n[A] = n_A \tag{4.35}$$

are all identical notations to denote the concentration of the neutral species A. The official SI unit[2] for concentration is moles per cubic decimeter (liter), mol dm^{-3}, where 1 mole of a substance contains 6.02×10^{23} particles, which is described by Avogadro's number $N_A = 6.02 \times 10^{23}\ \mathrm{mol}^{-1}$. For example, one mole of atomic oxygen would thus contain 6.02×10^{23} atoms. In ionospheric research, one is typically interested in the number of molecules or atoms per cubic centimeter or per cubic meter, thus, the number density of a given species is used as a proxy for the chemical concentration

$$n_A = \left[\frac{\#\ \text{molecules or}\ \#\ \text{atoms of species A}}{\mathrm{m}^3}\right], \tag{4.36}$$

where we used the hash sign "#" to mean "number of". The most common reaction in the ionosphere is the *two-body or bimolecular reaction* (4.28). Three-body reactions can also occur in the thermosphere, which has the form

$$\mathrm{A} + \mathrm{A} + \mathrm{C} \longrightarrow \mathrm{A}_2 + \mathrm{C} \tag{4.37}$$

where C is a third body.

For example, the principle source of the 557.7 nm green line emission in the MLT is caused by a three-body oxygen recombination, where the third body is atomic oxygen, which is during the recombination process partially excited. This reaction can be summarized by

$$\mathrm{O} + \mathrm{O} + \mathrm{O} \longrightarrow \mathrm{O}_2 + \mathrm{O}(^1\mathrm{S}), \tag{4.38}$$

where $\mathrm{O}(^1\mathrm{S})$ atoms either emit at 557.7 nm through a transition between the states $^1\mathrm{D}$ and $^1\mathrm{S}$ or are deactivated by oxygen molecules and do not emit radiation.

4.5.3 Reaction Rates

A *reaction rate* R expresses the time rate of change of chemical concentration of a reactant in a given reaction

$$\pm R = \frac{\mathrm{d}[A]}{\mathrm{d}t} \tag{4.39}$$

where $\pm$ denotes that the production $(+)$ or loss $(-)$ of the species A. The rate at which a chemical reaction occurs is extremely variable from one reaction to another reaction. Thus, every reaction has an associated *reaction rate*, which describes how fast a reaction occurs and is typically proportional to the concentration of the reactant. That means the more there is a specific species available, the higher is the occurrence chance of the reaction involving the species. The kinetic theory is used to obtain the general expression for the reaction rates. *A rate coefficient* k_r relates the reaction

[2]Unit adopted for international use under the Système International d'Unités. This sytem is used for all scientific purposes.

rate to the species concentration. The order of a rate coefficient is given by the order of the reaction. For example, unimolecular, bimolecular, and termolecular reactions have first-order k_1, second-order k_2, and third-order k_3 rate coefficients, respectively. The reaction rate R for these reactions can be expressed as

$$R = k_1[A] \qquad R = k_2[A][B] \qquad R = k_3[A][B][C], \tag{4.40}$$

where the rate coefficients k_1, k_2, and k_3 have the units s^{-1}, cm^3 $molecule^{-1}$ s^{-1}, cm^6 $molecule^{-2}$ s^{-1}, respectively. In general, a bimolecular reaction of the form

$$\mathrm{aA} + \mathrm{bB} \longrightarrow \mathrm{cC} + \mathrm{dD} \tag{4.41}$$

has the rate

$$R = k_r[A]^a[B]^b \tag{4.42}$$

where the rate coefficient k_r for the reaction between A and B is of order $a + b$. It is conclusive that when a species is treated as a reactant and product, such as the case for C in the three-body reaction (4.37), the rate of change of C is zero.

As different reactions have different reaction rates and in reality various chemical reactions occur in parallel to dynamical processes, it is obvious that chemistry and dynamics will compete in shaping the structure and evolution of the atmosphere and the actual system is very complex. In fact, all weather and climate models have to consider the coupling between dynamics and chemistry to adequately predict the future state of the atmosphere. You are probably familiar from weather forecast applications in smart phones that the predicted weather in the next seven days can still change significantly, which reflects the complexity of the chemical and dynamical processes in the atmosphere and thus how difficult it is to actually forecast the implications of multiple chemical and dynamical processes taking place simultaneously.

Both chemical and dynamical processes are temporally variable in the atmosphere. But how does one quantify which process is more important than the other? For this, it is convenient to evaluate the *chemical time constant* τ_c in comparison to the *dynamical time constant* τ_d. If the chemical time scale is larger than the dynamical time scale, i.e., $\tau_c > \tau_d$, then chemical processes are less important than the dynamical. If the chemical time scale is smaller than the dynamical time scale, then this means that chemical processes are occurring faster or more rapidly than dynamical processes, then the chemistry has to be taken into account as a controlling mechanism. In general, chemical processes tend to increase the spatial gradient in the concentration of species, while dynamical processes (dynamics) in form of small-scale eddy mixing can reduce the concentration gradient. One of the main physical mechanisms that contribute to mixing is the compositional effects induced by gravity waves (e.g., Garcia and Solomon 1985; Liu and Chester 2004; Walterscheid and Hickey 2012) and large-scale tides (Forbes et al. 1993). In a realistic atmosphere, both dynamics and chemistry should be considered in order to better model the general circulation.

The only first-order reaction that is significant for planetary ionospheres is the spontaneous deexcitation of a molecule, atom, or an ion. For example, the transition of an excited oxygen atom from ^{1}D state to the ground ^{3}P state represented by the following transition equation

$$O(^1D) \longrightarrow O(^3P) + h\nu \quad (\lambda = 630/636 \text{ nm}) \tag{4.43}$$

where energy is released in the form of a photon, $h\nu$, as the ^{1}D level is energetically higher than the ^{3}P, either at a wavelength of 630 or 636 nm, which corresponds to the oxygen red line. The energy of the photon is equal to $h\nu$, where h is the Planck constant. In this case, if a single atom undergoes this type of deexcitation, an energy of $E = hc/\lambda = 6.0263 \times 10^{-34}$ J s $\times$ 3×10^8 m s^{-1}/630×10^{-9} m = 2.87×10^{-19} J = 1.79 eV is released. In certain altitudes in the atmosphere, it is possible that the spontaneous deexcitation processes can occur faster than the transport processes, can thus not be neglected. The relevant rate equation is

$$\frac{d[O(^1D)]}{dt} = -k_1[O(^1D)] \tag{4.44}$$

Second-order reactions involving two different reactants are important in planetary ionospheres. For example, the ion-atom interchange reaction of the form

$$A^+ + B_2 \longrightarrow AB^+ + B \tag{4.45}$$

$$AB^+ + e^- \longrightarrow A + B \tag{4.46}$$

was presented in Eqs. (4.33)–(4.34). The associated rate equation is

$$\frac{d[O^+]}{dt} = \frac{d[N_2]}{dt} = -k_2[O^+][N_2], \tag{4.47}$$

where k_2 is a second-order rate coefficient and the contribution of transport processes to the loss process O^+ and N_2 are neglected in the above continuity equation.

The following third-order reaction

$$O + O + M \longrightarrow O_2 + M \tag{4.48}$$

would have the continuity equation

$$\frac{d[O]}{dt} = -2k_3[O]^2[M], \tag{4.49}$$

As the rate equations are the continuity equations for the chemical species and these equations are ordinary differential equations, they can be solved with the conventional method of harmonic wave solution method. For example, the rate equation presented in Eq. (4.44) for the deexcitation of the atomic oxygen in ^{1}D state has the solution

$$[\mathrm{O}(^1\mathrm{D})] = [\mathrm{O}(^1\mathrm{D})]_0 \exp[-k_1(t - t_0)] = [\mathrm{O}(^1\mathrm{D})]_0 \exp\left(-\frac{t - t_0}{\tau_1}\right), \tag{4.50}$$

where $\tau_1 = k_1^{-1}$ is the characteristic time constant, i.e., the time during which the initial concentration $[\mathrm{O}(^1\mathrm{D})]_0$ drops to e^{-1} of its initial value.

4.5.4 Chemical Continuity Equation

The chemical rate equations are in fact a form of the equation of continuity, which expresses the principle of conservation of mass in a small representative volume element. Each chemical compound or species s has its continuity equation. The principle of continuity had been discussed in the first volume of this monograph series where the continuity was given in the flux form by

$$\frac{\partial \rho_s}{\partial t} + \nabla \cdot (\rho_s \mathbf{u}_s) = 0 \tag{4.51}$$

which states that in the absence of sources and sinks the local time rate of change of the mass density of a species is balanced by the inward or outward flux of the species with respect to the infinitesimally small control volume. In other words, we can conclude from the continuity equation that if the mass density is increasing in a fixed volume element, i.e., $\partial \rho_s / \partial t > 0$, a convergence of mass is occurring (inward flux), i.e., $\nabla \cdot (\rho_s \mathbf{u}_s) < 0$. In the opposite scenario of mass density decrease, $\partial \rho_s / \partial t < 0$, in the system, divergence of mass, $\nabla \cdot (\rho_s \mathbf{u}_s) > 0$, must be occurring, again under the assumption of lack of sources and sinks. However, in a realistic atmosphere chemical production and loss occur, which we represent as a net source term

$$\Delta S_s = P_s - L_s \tag{4.52}$$

which is the difference between the production P_s and the loss L_s of the species s. The chemical continuity equation then becomes

$$\frac{\partial \rho_s}{\partial t} + \nabla \cdot (\rho_s \mathbf{u}_s) = \Delta S_s \tag{4.53}$$

Equivalently the continuity equation can be expressed in terms of the number density

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{u}_s) = \Delta S_s \tag{4.54}$$

4.6 Ionization Processes and Solar Radiation

We have briefly discussed in Sect. 2.8 the nature of the electromagnetic radiation. Radiation can be absorbed, emitted, and scattered. Our Sun emits a broad range of wavelengths (energies). For example, Fig. 2.6 presented the solar irradiance spectrum measured by the SORCE spacecraft, which can provide precise measurements for the wavelengths from 1 to 2000 nm. How much energy will be absorbed by a planetary atmosphere depends on the neutral composition of the atmosphere. Not all wavelengths contribute to ionization. Different neutral species require different amount of minimum ionization energy to be ionized. For example 13.62 eV, corresponding to an electromagnetic wavelength of 91 nm, is required to free an electron from the oxygen atom O. It is 12 eV (102 nm) for the oxygen molecule O_2, 13.77 eV (90 nm) for CO_2, and 9.26 eV (133.8 nm) for NO. Various minimum ionization threshold energies relevant to planetary atmospheres are summarized in Table 4.2.

When radiation is absorbed from Sun within our atmosphere, molecules are dissociated and free electrons form. The process of the formation of a free electron-ion pair is called *ionization*. The vertical ionization profile of a planetary atmosphere can be considered by a simple model: At the top of the atmosphere, the mass density is low, so the rate of ionization is relatively small. However, deeper in the atmosphere there are more neutral particles that can be ionized than at the top of the atmosphere and the rate of ionization increases. At even deeper layers in the atmosphere the strength of the radiation decreases so that the ionization eventually decreases. Thus, there is a height in the upper atmosphere where ideal conditions are present for the maximum amount of ionization to occur as illustrated with horizontal red dashed lines in Fig. 4.4, where idealized neutral density and ionization rates are sketched as well.

In planetary atmospheres, the radiative transfer calculations of the solar energy deposition is simple because the dominant radiative process is absorption. Specifically, we shall next focus on the process of solar radiation absorption, which is the primary source of ionization in a planetary atmosphere. The analysis of absorption of radiation deals with the process of how the incoming radiation intensity varies as it penetrates downward in the atmosphere. The radiant energy dE_λ in a specific wavelength interval $(\lambda, \lambda + d\lambda)$ is commonly understood as the intensity I_λ asso-

Table 4.2 Ionization energies and wavelengths of some important neutral species in planetary atmospheres

Species		Ionization wavelength [nm]	Ionization energy [eV]
NO	Nitric oxide	133.8	9.26
O_2	Molecular oxygen	102.6	12.1
O	Atomic oxygen	91.1	13.62
CO_2	Carbon dioxide	90	13.77
N_2	Molecular nitrogen	79.6	15.58

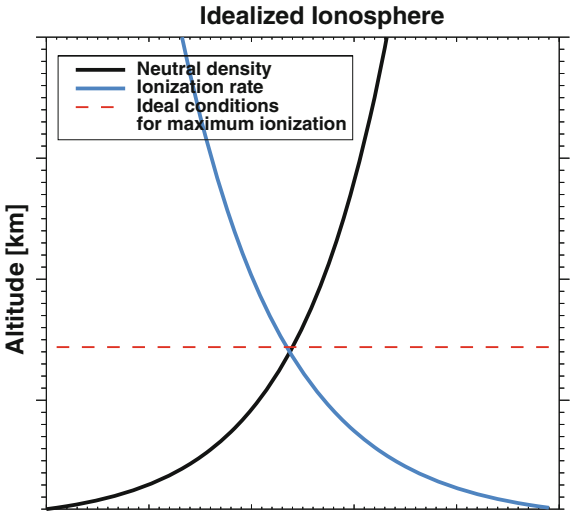

Fig. 4.4 Illustration of an idealized planetary ionosphere in which the ionization rate drops exponentially deeper in the atmosphere and the neutral mass density drops exponentially with increasing altitude. At there location where both curves intersect marked with *red dashed lines* the optimal conditions for maximum ionization is expected

ciated with that wavelength interval. In reality, this intensity is expected to have spatiotemporal variations. A radiation field that is independent of direction at a point is said to be isotropic at that point. A field that is the same at all points and in all directions is called homogeneous and isotropic.

4.7 Chapman Layers

In order to state the necessary mathematical treatment, it is important to discuss various simplifications and assumptions that are made in the absorption problem. In a planetary atmosphere context, it is assumed that the atmosphere is horizontally stratified in parallel planes. There is a single absorbing species, which decreases exponentially with increasing altitude with a constant characteristic length scale. The simplification is illustrated in Fig. 4.5, where χ is the solar zenith angle, i.e., the angle that the incoming radiation makes with the vertical, and z is the height. The incoming radiation is monochromatic (consisting out of one wavelength λ). Let σ_a denote the absorption cross section, $n_n(z)$ the neutral density, s_λ distance along the path of the penetrating photons, and $I(s_\lambda)$ the photon flux (intensity) along the distance of the incoming photons. As the photons are absorbed by the neutrals when they penetrate downward in the atmosphere, the intensity decreases by an amount of $\mathrm{d}I(s_\lambda)$ over a distance of s_λ and is given by

$$dI(s_\lambda) = -I(s_\lambda)\, n_n(z)\, \sigma_a\, \mathrm{d}s_\lambda, \tag{4.55}$$

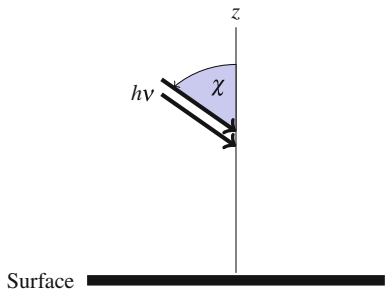

Fig. 4.5 Illustration of the incoming solar radiation $h\nu$ at the top of the atmosphere at a zenith angle χ, the angle with the vertical, and the height z

where the minus sign indicates a decrease. Within the plane stratified atmosphere assumption, the distance the photon travels can be expressed as

$$\mathrm{d}s_\lambda = -dz \sec\chi, \tag{4.56}$$

where $\sec\chi = (\cos\chi)^{-1}$. If we integrate over all the incrementally small absorbed intensity elements, we will get the total absorbed intensity as integrated from a reference level z to infinity

$$\int_z^\infty \mathrm{d}I(s_\lambda) = \int_z^\infty I(s_\lambda) n_n(z)\sigma_a \sec\chi \,\mathrm{d}z \tag{4.57}$$

$$\Rightarrow \int_z^\infty \frac{\mathrm{d}I(s_\lambda)}{I(s_\lambda)} = \int_z^\infty n_n(z)\sigma_a \sec\chi \,\mathrm{d}z \tag{4.58}$$

$$\Rightarrow [\ln I]_z^\infty = \ln\left(\frac{I_\infty}{I_z}\right) = \int_z^\infty n_n(z)\sigma_a \sec\chi \,\mathrm{d}z, \tag{4.59}$$

where I_∞ is the unattenuated photon flux at the top of the atmosphere and the variation of the number density of the neutrals with altitude is approximated by

$$n_n(z) = n_n(z_0)\exp\left(-\frac{z - z_0}{H}\right), \tag{4.60}$$

where $n_n(z_0)$ is the neutral number density at a reference altitude z_0 lower in the atmosphere and H is the scale height, which for simplification is assumed to be constant. So, how do we obtain the photon flux at an arbitrary altitude z? From (4.59) we get

$$I(z, \chi) = I_z = I_\infty \exp[-H\, n(z)\sigma_a \sec\chi]. \tag{4.61}$$

In reality, this expression has some limitations. Namely, in a realistic world, the photon flux and the cross section are a function of wavelength. There are more than one absorbing atmospheric species with its own scale height variations and the planet is not flat. So, we should start with an expression that is more general than Eq. (4.55) for the infinitesimally small decrease in intensity as the photons

enter the atmosphere, taking into account for the presence of different species whose contributions to intensity absorption depend on the associated wavelength-dependent species absorption cross section $\sigma_{a,s}(\lambda)$:

$$\mathrm{d}I(z, \lambda, \chi) = -\sum_s n_s(z)\, \sigma_{a,s}(\lambda) I(z, \lambda) \mathrm{d}s_\lambda, \tag{4.62}$$

where summation is performed over the different species s, each species has its own cross-section and the photon intensity depends on the wavelength as well as the altitude. Integrating this more general form (4.62) of the intensity decrease along the optical path yields

$$I(z, \lambda, \chi) = I_\infty(\lambda) \exp\left(-\int_\infty^z \sum_s n_s(z) \sigma_{a,s}(\lambda) \mathrm{d}s_\lambda \right) \tag{4.63}$$

$$= I_\infty(\lambda) \exp[-\tau(z, \lambda, \chi)], \tag{4.64}$$

where we have defined the *optical thickness* (or *optical depth*) τ as

$$\tau(z, \lambda, \chi) \equiv \int_\infty^z \sum_s n_s(z) \sigma_{a,s}(\lambda) \mathrm{d}s_\lambda. \tag{4.65}$$

Let us consider two idealized scenarios: (1) an absolutely transparent atmosphere would have zero optical thickness (or opacity), which would mean no absorption takes place within the atmosphere and thus the intensity at a given altitude would be equal to the intensity at the top, i.e., $I_z = I_\infty$. In an opaque atmosphere with non-zero optical thickness ($\tau > 0$), we have a reduction of intensity at the layers lower down in the atmosphere compared with the initial radiation at the top.

In a planetary atmosphere, the altitude variations of the density of the neutral species depend on the temperature variations and the species scale height as

$$n_s(z) = n_s(z_0) \frac{T_s(z_0)}{T_s(z)} \exp\left(-\int_{z_0}^z \frac{\mathrm{d}z'}{H_s(z)} \right), \tag{4.66}$$

where the species scale height is given by

$$H_s(z) \equiv \frac{kT_s(z)}{m_s\, g(z)}, \tag{4.67}$$

where m_s is the molecular mass of the species and $g(z)$ is the height-dependent gravitational acceleration. In the approximation of isothermal atmosphere with constant gravity the species scale height $H_s = kT_s/m_s g$ is constant and $T_s(z) = T_s(z_0)$, hence the number density variation takes the simplified form

$$n_s(z) = n_s(z_0) \exp\left[-\left(\frac{z - z_0}{H_s}\right)\right], \tag{4.68}$$

which is similar to (4.60). Integration of the density from some reference level z_0 to infinity yields the vertical column density

$$\int_{z_0}^{\infty} n_s(z)\mathrm{d}z = -H_s\, n_s(z)\Big|_{z_0}^{\infty} = n_s(z_0) H_s \tag{4.69}$$

The above simplified discussion of absorption of photons within the atmosphere helps illustrate a practical representation of the ion production process in the ionosphere as was initially discussed in the work by Chapman (1931). As emphasized before the basic mechanism that is responsible for the formation of the ionosphere is photoionization, which leads to a free electron-ion pair. If the incoming radiation provides more energy than required for the ionization of a given neutral species than the excess energy is taken by the free electron as a kinetic energy and/or the resulting ion is excited. The rate of production of electron-ion pair can be represented by the Chapman production function P_c as

$$P_c(z, \chi) = I(z, \chi)\eta\sigma_a\, n(z), \tag{4.70}$$

where the intensity of photon radiation at a reference level, $I(z, \chi)$ is given by Equation (4.61) for the simplified case of single absorbing species and η is a probability of photon absorption. Rewriting this expression yields

$$P_c(z, \chi) = I_\infty \exp[-H\, n(z)\sigma_a \sec\chi]\, \eta\, \sigma_a\, n(z). \tag{4.71}$$

The maximum rate of ionization at an altitude z_m occurs when $\partial P_c/\partial z = 0$. For this condition, we get the location of the maximum ionization production

$$P(z = z_m, \chi) = \frac{I_\infty\, \eta \cos\chi}{H\, \exp(1)} \tag{4.72}$$

Note that the different atmospheric species have different ionization threshold potential (see Table 4.2). This threshold has to be overcome in order to generate an electron-ion pair.

4.8 Earth's Ionosphere

Just like the way Earth's neutral atmosphere has a distinct vertical structure, Earth's ionosphere has a distinct vertical structure as well. Its overall vertical structure results from the altitude variation of ionizable species and of the ionization rate. The ionosphere is composed of the D, E, and F regions, where D region is the lower-

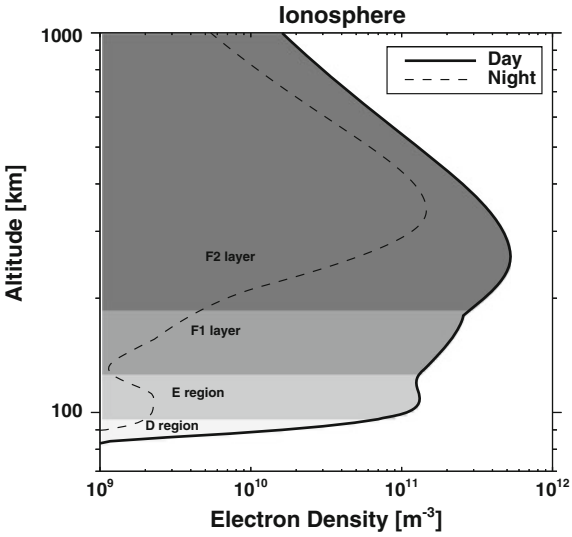

Fig. 4.6 IRI 2012 model

most region and F region is the top of the ionosphere. Additionally, the F region can be described in terms of the F_1 and F_2 layers. All known planetary ionospheres demonstrate a vertical structure, which is generally characterized according to the vertical profile of the electron density (e.g., Benna et al. 2015; Withers et al. 2015; Witasse et al. 2008; Ratcliffe 1972). Specifically, this division arises from the successive peaks of electron density n_e. Figure 4.6 presents the altitude profile of the nigthtime (dashed) and daytime (solid thick) electron density in the terrestrial ionosphere based on the International Reference Ionosphere 2012 ((IRI), Bilitza et al. 2014).

An earlier theory of the formation of ionospheric layers based on idealized assumptions was already presented in the 1930s by Chapman (1931) as discussed in the work by Yonezawa (1966). Since then a great deal of progress has been made in the detailed understanding of various ionospheric processes.

In a plasma, the index of refraction

$$n_r = \frac{c}{v}, \tag{4.73}$$

i.e., the ratio of the speed of light in vacuum c to the propagation speed v of the signal, is inversely proportional to the wave frequency. This property implies that the higher the frequency of a given wave, the smaller is n_r and thus the larger is the propagation speed. From the definition of the refractive index and noting that the frequency of a given wave does not change in a refractive medium, we get

$$n_r = \frac{\lambda_0}{\lambda} = \frac{k}{k_0}, \tag{4.74}$$

where λ_0 and k_0 are the wavelength and wave number in vacuum, respectively, related by $k_0 = 2\pi/\lambda_0$. For $n_r > 1$, $\lambda < \lambda_0$ meaning that the wavelength of a wave signal in a refractive medium is shorter than its wavelength in vacuum. By definition, $n = 1$ is the index of refraction in vacuum. Then we have $\lambda = \lambda_0$. Any other medium has a refractive index larger than unity. For example, air has an index of 1.0003 and water has $n = 1.3335$. Assuming that the ionosphere is horizontally stratified with an upward increasing number of electrons, a plane wave incident vertically will be reflected to the ground if $n_r = 0$. A signal propagating through the ionosphere experiences in general amplitude descrease with distance as $\exp(\kappa z)$, where κ is an absorption coefficient. The most generic form of the refractive index can be expressed as a complex number

$$n_r = a_r - i b_r, \tag{4.75}$$

where the real part a_r expresses the wave velocity and the imaginary part b_r represents the absorption. As the ionosphere is a partially ionized medium it can be considered a partial conductor. Thus, the amount of absorption in the ionosphere is related to the electric conductivity.

Historically, the refractive properties of the ionospheric plasma have been exploited in order to characterize the structure of the ionosphere (Ratcliffe and Weekes 1960; Schunk and Nagy 2009). Specifically, the standard method of determining the vertical structure has been the radio sounding technique. In this technique, the time delay τ between the transmitted and the received radio signal (echo) is used to determine the height of the reflecting layer, which is referred to as the virtual height h' given by (Elias et al. 2017)

$$h' = \frac{c\,\tau}{2}, \tag{4.76}$$

where c is the speed of light. Typically, as the transmitted frequency of the wave is increased the respective altitude of reflection increases until a so-called *critical frequency* or the *penetration frequency* f_0 is reached at which there are just sufficient electrons for the reflection to take place. In other words, this critical frequency is given by the maximum plasma frequency of a given layer. The critical frequency is proportional to the electron density and is related to a virtual height. Each ionospheric layer is characterized by its own critical frequency, such as, f_0E, f_0F_1, and f_0F_2. Above the critical frequency no reflection takes place. If a signal frequency is larger than the critical frequency, the signal propagates through the layer. Therefore, if the signal frequency emitted by a ground-based source is larger than the maximum critical frequency for the whole ionosphere, then information can be received by space-borne instruments.

An *ionosonde* instrument uses the reflective properties of the ionosphere in order to determine various ionospheric parameters. An ionosonde recording is called an *ionogram*, which displays a unique relationship between the frequency of the emitted signal and the virtual height h'. A sample ionogram is seen in Fig. 4.7 adopted from the work by Reinisch (2000). The structure of the echo is presented in a

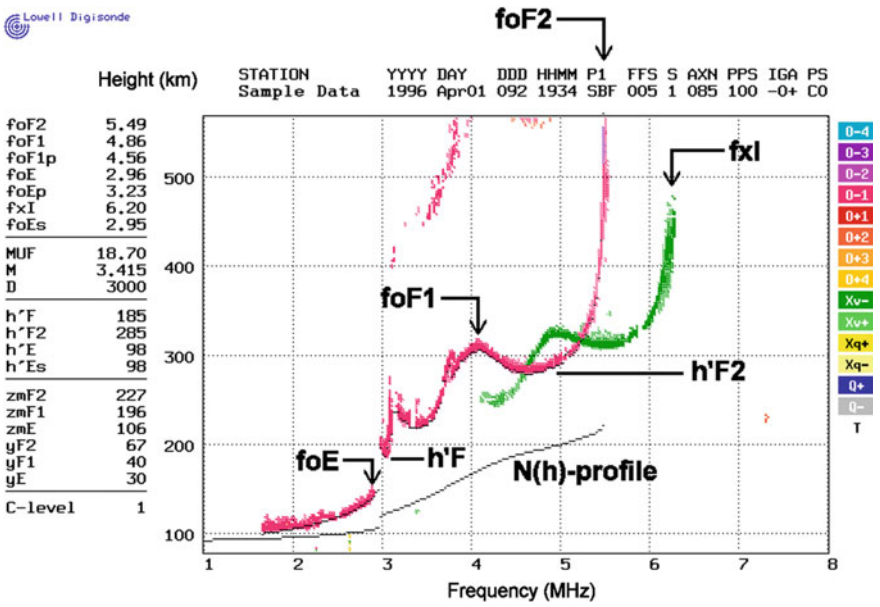

Fig. 4.7 A sample ionogram for midday recorded at Millstone Hill, MA displaying the different ionogram outputs: O (*red*) and X (*green*) echo traces and illustrating the different ionospheric layers. Some of the auto scaled characteristics are listed on the *left* (Reinisch 2000, Fig. 1)

height-frequency diagram, where the critical frequencies f_0E, f_0F_1, and f_0F_2 and the virtual heights $h'F$ and $h'F_2$ are retrieved. The derived electron density profile is overplotted, denoted by $N(h)$ in the figure. Modern-day ionosondes typically scan a frequency range of 1–30 MHz, corresponding to wavelengths ranging from 10 to 300 m.

Processes influencing the formation and evolution of a planetary ionosphere are controlled by *photochemical* processes and *transport* processes depending on the height. Additionally, the ionosphere is influenced by external forces from below (Laštovička 2006; Pancheva et al. 2012; Yiğit et al. 2016) and from above (Laštovička 2009; Yiğit and Ridley 2011; Yiğit et al. 2012; Gong et al. 2013). Photochemical processes involve the production and destruction of ionization while transport represents how ionization moves or gets redistributed.

When characterizing the different ionospheric regions in detail, three important physical aspects should first be considered

1. The solar spectrum deposits its energy at various heights depending on the absorption characteristics of the atmosphere, which in turn depends on the neutral (ionizable) species;
2. Recombination depends on the density; and
3. Atmospheric composition changes with height.

The interplay of these properties lead to the formation of ionospheric regions with different governing processes and varying degree of complexity.

If both photochemical and dynamical (i.e., transport) processes are considered, the tendency of electron concentration will result from a superposition of production P, loss L, and transport

$$\frac{\partial n_e}{\partial t} = P - L - \nabla \cdot (n_e \mathbf{u}_e), \tag{4.77}$$

where the transport term

$$\nabla \cdot (n_e \mathbf{u_e}) = n_e(\nabla \cdot \mathbf{u}_e) + \mathbf{u}_e \cdot (\nabla n_e), \tag{4.78}$$

depends on the divergence of the electron drift velocity $\mathbf{u}_e$, $\nabla \cdot \mathbf{u}_e$, and on the concentration gradient, ∇n_e. Note that the structure of the electron concentration tendency equation (4.77) is identical to the chemical continuity equation (4.53) or the number density equation (4.54). Equation (4.77) applies equally to ions.

Various approximations can be made to the electron concentration equation. In Earth's ionosphere below 200 km, photochemical equilibrium assumption is under certain circumstances a good approximation in the D, E, and F_1 regions by day. In this case, the electron tendency controlled solely by the production and loss of electrons. If the loss process is occurring much faster than the time rate of change of the electrons, than *photochemical equilibrium* applies, in which production term balances the loss term, $P = L$. Above about 250 km in the F_2 layer photochemical processes are less important than transport, thus the $\nabla \cdot (n_e \mathbf{u_e})$ term is the dominant term.

The solar energy is attenuated while it passes through the atmosphere because of ionization, dissociation, and excitation of atmospheric constituents. Photoionization is the predominant production process, while recombination is an important loss process in Earth's ionosphere. *Photoionization* occurs via the absorption of a photon

$$\mathrm{A} \xrightarrow{h\nu} \mathrm{A}^+ + \mathrm{e}^-, \tag{4.79}$$

where the photon energy $h\nu$ must be larger than the ionization potential of the species A. The ionization potential is the minimum amount of energy that is necessary to remove an electron from an atom or molecule to infinity. Here we are mainly concerned with the outermost electrons, that is, the electrons which are in the outermost orbit or shell of an atom. These electrons are sometimes referred to as valence electrons.

Radiative recombination is the reaction of an ion with a free electron

$$\mathrm{A}^+ + \mathrm{e}^- \longrightarrow \mathrm{A}^* + \mathrm{h}\nu, \tag{4.80}$$

where A^* indicates that the species A may be in an excited state. The predominant loss process is the electron-ion dissociative recombination

$$AB^+ + e^- \longrightarrow AB \longrightarrow A^* + B^* + h\nu, \tag{4.81}$$

A^* and B^* correspond to the excited states of the constituents A and B, respectively. This process is of particular importance in Earth's, Venus', and Mars' ionospheres for the loss of the molecular oxygen ion

$$O_2^+ + e^- \longrightarrow O^* + O^* + h\nu, \tag{4.82}$$

where O^* indicates that the product can be in an excited state of the atomic oxygen, such as $O(^1D)$ and $O(^1S)$.

Essentially, the photoionization produces a free electron-ion pair. The excess energy goes to the kinetic energy of the electron and excitation of the produced ion. A convenient description of the photoionization process is given by the *Chapman production function*, P_c, presented in Eq. (4.70). Let us next discuss some of the basic features of Earth's ionosphere, which is the most well-studied planetary ionosphere and many fundamental processes can be generalized to other ionospheres.

4.8.1 D Region

The *D* region is the lowest lying region of the ionosphere, extending from $\sim$ 60 to $\sim$90–95 km. It is a region of weakly ionized plasma with densities in the order of 10^6 m^{-3}, typically disappears at night, and coexists with a relatively denser mesosphere and lower thermosphere (MLT) that is controlled by neutral dynamics and internal waves. The ionization sources are very energetic charged particles, such as the Galactic Cosmic Radiation (GCR), the most energetic part of the solar irradiation, which is in the X-ray spectrum range with wavelengths less than 10 nm and energies in the order of 100 keV (Fig. 2.5), EUV radiation less than about 110 nm, and hydrogen Lyman-α emission line (121.6 nm). The minor constituent NO is strongly ionized by the strong portion of the ultraviolet radiation.

4.8.2 E Region

The *E* region is the first ionospheric region that was discovered. It extends from 90–95 to $\sim$130 km. Ionization of atomic and molecular oxygen contributes to the formation of E region by high-energetic X-ray radiation and Lyman β radiation (102.6) nm. The predominant ions are O_2^+ and NO^+ and the typical plasma density is in the order of 10^{11} m^{-3}. Additionally, meteor ablation between about 85 and 130 km leads to the

formation of metallic ions, such as Fe^+, Si^+, and Mg^+. The E region coincides with the dynamo layer and is the region where large electrical conductivities are found.

4.8.3 F Region

The F region ionosphere plays a key part in ionospheric research owing to its complexity and anomalous behavior. It can be split into an F_1 layer extending from $\sim$ 130 to $\sim$ 180 km and the F_2 layer, extending from 180 km upwards. In terms of the chemistry and physics there are fundamental differences between the F_1 and F_2 layers. The F_1 is overall in photochemical equilibrium and it disappears at night, while in the F_2 layer dynamical and transport processes are important and it is thus not in photochemical equilibrium and cannot be described by a Chapman layer approximation. Also, the F_2 layer exists during the day and night. In particular above 250 km transport is very important, where the chemical loss rate of ions is comparable to their diffusion rate through the neutral gas. Hence, transport directly influences electron and ions densities. The F_2 plasma density peak occurs as a result of competition of chemical loss (recombination) and diffusive transport. The height of the peak plasma density, $h_m F_2$, is highly variable: 200–450 km, depending on solar and geomagnetic activities. The dominant species in F_2 layer is O^+, while in the F_1 layer O_2^+ and N_2^+ are prevalent. The atomic oxygen red line emissions take place in the F-region, in particular, in the F_2 layer at night. F region is also the most anomalous ionospheric region, demonstrating e.g., Spread-F and annual asymmetry phenomena. It is strongly coupled to magnetospheric activity via current systems and convection electric fields at high-latitudes.

4.8.4 Dynamo Layer

In the mesosphere and lower thermosphere (80–130 km), coinciding with the D and E region ionosphere in altitude, mutual interaction of neutrals and plasma lead to formation of a dynamo layer, where neutral dynamics, mainly dominated by tidal motions and gravity wave saturation, can lead to dynamically induced electric currents in the presence of a magnetic field. The plasma and neutral parameters that shape the altitude variation of this layer are the plasma gyrofrequencies and collisions between plasma and neutrals. As collision frequencies depend on the mass density of the species involved in collisions, it varies in general much strongly with altitude than the gyrofrequency variation with altitude.

The gyrofrequency of electrons are much larger than ions for a given magnetic field due to their much smaller mass. Within the dynamo region ions are tightly coupled to the neutrals as $\nu_{in} > \omega_{c_i}$, while for the electrons we have $\nu_{en} < \omega_{c_e}$ and thus electron motion is governed primarily by the $\mathbf{E} \times \mathbf{B}$ drifts, driving the Hall current associated

with the Hall conductivity. At the top of the dynamo layer $\nu_{in} \approx \omega_{c_i}$, where ions lead to a Pederson current.

Overall, in the low-latitude to middle-latitude dynamo region, the generalized Ohm's law describes the relationship between electric fields and conductivities as (Vichare et al. 2012)

$$\mathbf{j} = \sigma_p(\mathbf{E}_\perp + \mathbf{u} \times \mathbf{B}) + \sigma_H \hat{\mathbf{b}} \times (\mathbf{E}_\perp + \mathbf{u} \times \mathbf{B}) + \sigma_\parallel \mathbf{E}_\parallel, \tag{4.83}$$

where $\mathbf{u}$ is the neutral drift velocity, $\mathbf{E}_\parallel$ and $\mathbf{E}_\perp$ are the components of the electric field parallel and perpendicular to the magnetic field, respectively, and $\hat{\mathbf{b}} = \mathbf{B}/|\mathbf{B}| = \mathbf{B}/B$ is the unit vector along the magnetic field.

4.9 Mars Ionosphere

Our closest neighbor Mars, the fourth planet from Sun, has fascinated astronomers since the early ages because of its spectacular similarities and but also interesting differences to Earth.

4.9.1 General Characteristics

Mars is also a terrestrial planet with a prominent atmosphere. Its radius is approximately the half of Earth's ($r_m/r_e = 0.532$); with a mean distance of 1.523 AU from Sun, it has a mean orbital period of 686.9 sols; its axial tilt of 25.19° and rotational period (one rotation around its own axis, i.e., 1 sol) of 24.6 h, which is only about 2.5% longer than an Earth day, are very similar to those of Earth. However, Mars's orbit around Sun is very eccentric, producing substantial seasonal differences. Its distance from Sun varies from ~1.34 AU during winter solstice ($L_s = 270°$) to ~1.65 AU during summer solstice ($L_s = 90°$) in the northern hemisphere. Thus, this orbital configuration implies that the seasons in the different hemispheres occur with varying intensity. For example, Mars is some 45 million kilometers closer to Sun during southern hemisphere summer than northern hemisphere summer, which are called the *perihelion*, the closest point to Sun, and the *aphelion*, the furthest point of the orbit from Sun. If you realize that the mean distance of Moon from Earth is about 380 thousand kilometers, this variation of the orbital distance is pretty remarkable. How elliptical the orbit of a planet around Sun is quantified by the orbital eccentricity parameter, which is $e = 0.0934$ for Mars, where the eccentricity e is evaluated in terms of the semimajor axis a and semiminor axis b as $e = \sqrt{1 - (b^2/a^2)}$. In comparision, Earth's current orbital eccentricity is 0.0167.

4.9.2 Formation of the Ionosphere

Mars' atmosphere, extending from its surface up to about 300 km, is composed of 95% CO_2. Thus, at high-altitudes, CO_2 is the main species which can be photoionized. This process is analogous with the photoionization of O_2 and N_2 in Earth's thermosphere. The atomic oxygen is a minor species. In Martian models, due to insufficient knowledge of the distribution and variability of oxygen, one has often used oxygen profiles based on the one-dimensional photochemical model of Nair et al. (1994). On Mars, ion production thus occurs by the ionization of CO_2 molecules by solar photons of $\lambda \leq 90$ nm and of O atoms by photons of $\lambda \leq 91.1$ nm:

$$CO_2 \xrightarrow{h\nu} CO_2^+ + e^-, \tag{4.84}$$

and

$$O \xrightarrow{h\nu} O^+ + e^-, \tag{4.85}$$

However, a major ion is the molecular oxygen ion O_2^+, which is quickly formed via atom-ion exchange between the atomic oxygen and the carbon dioxide ion (CO_2^+) as

$$O + CO_2^+ \longrightarrow O_2^+ + CO \tag{4.86}$$

or by charge transfer

$$O + CO_2^+ \longrightarrow O^+ + CO_2 \tag{4.87}$$

$$O^+ + CO_2 \longrightarrow O_2^+ + CO \tag{4.88}$$

The Martian ionosphere has a distinct vertical structure, which can be retrieved for example from satellites. Rishbeth and Mendillo (2004) have analyzed Mars Global Surveyor electron density data and evaluated 17 daily mean electron density profiles for March 1999. Their results are shown in Fig. 4.8. There is a certain degree of day-to-day variability but the overall structure of the profiles is robust. The lower layer

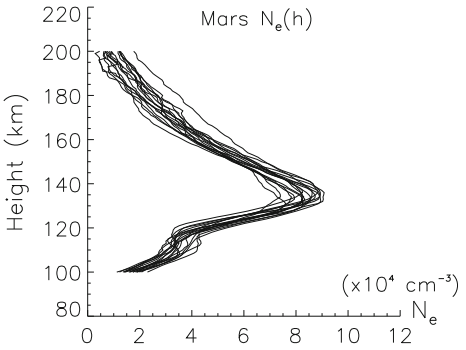

Fig. 4.8 Altitude variations of the daily mean electron density profiles [cm^{-3}] taken in March 1999 from 9–27 March 1999 at latitudes 69−73° N around 0400 local time at solar zenith angle $\chi = 76.5$–78° (Rishbeth and Mendillo 2004)

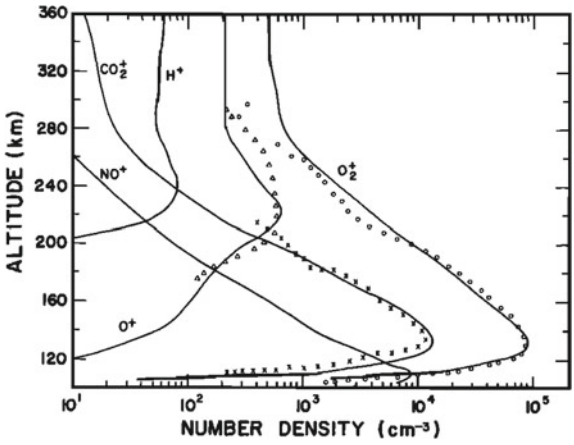

Fig. 4.9 Altitude variations of the number densities [cm^{-3}] of ion species, NO^+, CO_2^+, O^+, O_2^+, and H^+, in the Martian ionosphere calculated with a theoretical ionosphere model, coupling plasma conservation equations and comparison with Viking 1 and 2 lander observations for CO_2^+ (*crosses*), O^+ (*triangles*), and O_2^+ (*circles*). (Chen et al. 1978, Fig. 6)

called the "M1" layer peaks at a height of about 110 km above the Martian surface. The main layer called "M2" peaks at about 135 km. Rishbeth and Mendillo (2004) suggest that the M1 and M2 layers are comparable to the terrestrial E and F_1 layers.

Earlier measurements of the Martian ionosphere have been conducted with the Viking mission and interpreted with theoretical ionosphere models. Using such an ionosphere model, Chen et al. (1978) have calculated ion density and the temperatures of ions and electrons by solving the coupled continuity, momentum, and energy equations. Their results are seen in Fig. 4.9, where the vertical profiles of NO^+, CO_2^+, O^+, O_2^+, and H^+ are shown, specifically, comparing with Viking 1 & 2 lander observations of CO_2^+ (crosses), O^+ (triangles), and O_2^+ (circles), which suggested an overall good agreement.

Recently, NASA's MAVEN mission detected with the Neutral Gas and Ion Mass Spectrometer (NGIMS) instrument 22 ion species and measured the associated vertical profiles between 150 and 500 km (Benna et al. 2015), which provides further details of the Martian topside ionosphere. They showed that below 300 km, O_2^+ is the dominant ion species, however, O^+ number density peaks around 300 km ($n_{O^+} = 10^3$ cm^{-3}) becomes comparably abundant above this altitude as it was observed that O^+ and O_2^+ densities were approximately equal. Figure 4.10 presents the altitude profiles of the averaged density of the ion species as observed by MAVEN/NGIMS. Both MGS and MAVEN observations show that the peak plasma density in the Martian ionosphere is in the order of 10^{10} m^{-3}, which is about 100 times less than the peak electron density in Earth's ionosphere.

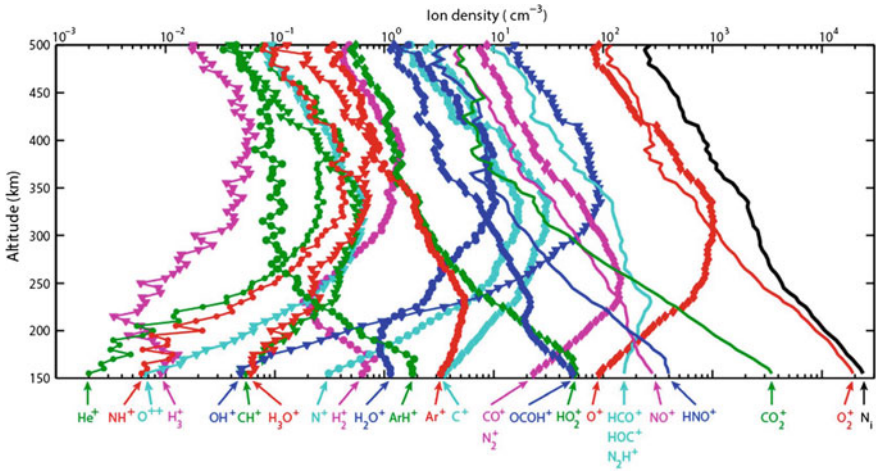

Fig. 4.10 Altitude profiles of the averaged density of ionospheric ions measured by NGIMS at SZA = 60 at altitudes between 150 and 500 km. (Benna et al. 2015)

Comparison with the measurements of the Viking mission, MAVEN observations of neutral and ion species (O^+, O_2^+, CO_2^+,) have demonstrated much larger densities (Withers et al. 2015). Previous modeling studies have demonstrated some degree of discrepancy in terms of which is the dominant ion species in the topside ionosphere. Seasonal effects can play a role in this context, which most observations and models have yet not considered.

4.10 Space Weather

Space weather is understood as the effects of Sun on Earth's magnetic field and atmospheric environment. As the magnetosphere is coupled to the ionosphere, and the ionosphere coexists with the thermosphere, any solar disturbances can be transferred to Earth's atmosphere. Geomagnetic storms are clear manifestations of space weather effects. During such storms sudden disturbances occur in the intrinsic magnetic field with a resultant coupling to the ionosphere-thermosphere. Figure 4.11 illustrates Earth's Sun-atmosphere environment coupled to the technological infrastructure, such as satellites, electric power grids, aviation, and navigation. The causes and effects of space weather events should be better understood in order to better forecast these events and take the necessary measures. For example, space radiation in the form of high-energy protons originating at Sun can affect critical computer systems during rocket launches. Efffects of space weather will be further discussed in the next Chap. 5.5.

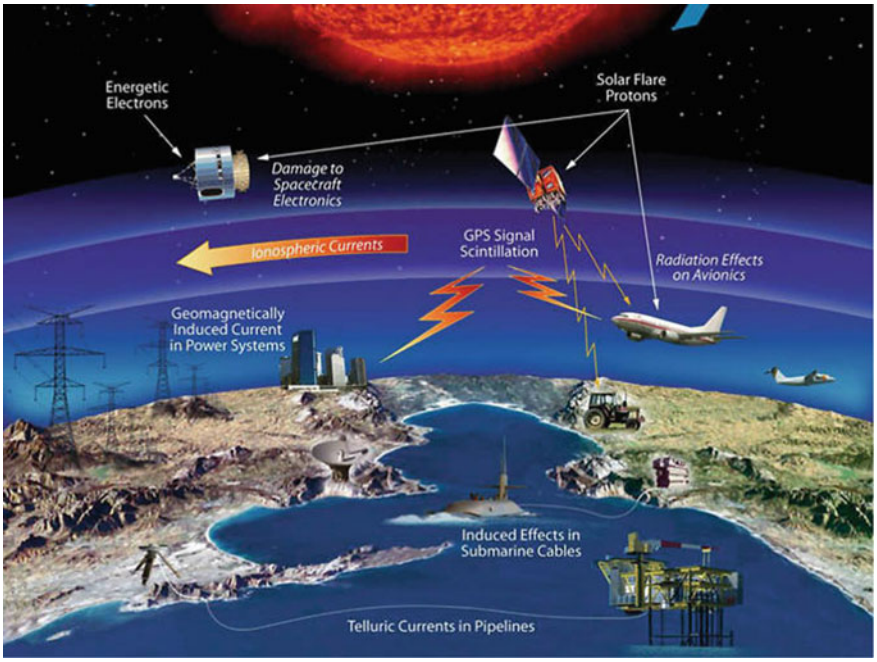

Fig. 4.11 Technological infrastructure affected by space weather events include satellites, aircraft, and power grids. Credit: NASA

4.11 A Summary of the Key Properties of Planetary Ionospheres

So far I have discussed some of the physical processes that shape the morphology of planetary ionospheres, focusing on the terrestrial ionosphere. I have also presented a brief discussion of the Martian ionosphere. As the ionosphere coexists with the neutral atmosphere, there is a significant amount of coupling between the neutrals and ions, depending on the degree of ionization and their differential motion. In Earth's upper atmosphere, the ionosphere is less dense than the neutral atmosphere ($n_e < n_n$), however, it extends over a much larger altitude than the thermosphere. Some general properties of planetary ionospheres can be summarized as follows:

Formation Above a certain altitude in the upper atmosphere, for example $z > 60$ km for Earth, photoionization becomes significant and plays the major role in the formation of the ionosphere, which to a large extent overlaps with the thermosphere, the hottest region of a planetary atmosphere.

Characterization The ionosphere is characterized in terms of the vertical variation of the plasma density. On Earth, we have the D-, E-, and F-region, where the latter

contains the largest amount of plasma. On Mars, the ionosphere is described in terms of the M1 and M2 layers, which are comparable to the terrestrial E and F_1 layers.

Collisions The ionosphere is a multi-species plasma. Electrons, ions and neutrals interact with each other via collisions. These interactions are expressed by the associated collision frequencies ν_{ij}, expressing the collisions of the species i with the species j. The collisions play an important role in the energy and momentum budget of the thermosphere-ionosphere.

Chemistry Ion chemistry is an important process in the ionosphere. Photoionization produces a free electron-ion pair. Thus, positive ions are the dominant plasma species, such as O^+ and N_2^+ in Earth's ionosphere; O^+ and CO_2^+ in Venus' ionosphere; and O_2^+ and O^+ in Mars' ionosphere. All ionospheric species demonstrate distinct variations with altitude, depending on their scale height. Negative ions can form as well, for example, in Earth's D-region ionosphere, where the neutral density is relatively larger. Charge exchange, recombination, and plasma excitation influence further the production and loss processes.

External influences The ionosphere is influenced, by meteorological processes (coupling from below) and by space weather effects (coupling from above). Meteorological effects encompass internal wave propagation and dissipation. Space weather effects describe the impact of Sun and magnetic processes.

Variability Due to the interplay of internal processes (e.g., weather, chemistry, collisions, transport, etc.) and external influences planetary ionospheres demonstrate a great deal of variability. This variability manifests itself often as wave-like variations in field variables. Ionospheric electron densities can demonstrate variability on different spatiotemporal scales.

Next chapter will discuss the dynamics of the upper atmosphere, focusing on the effects of meteorological processes (coupling from below) and space weather effects (coupling from above).

References

Appleton EV (1932) Wireless studies of the ionosphere. Inst Electr Eng-Proc Wirel Sect Inst 7(21):257–265

Benna M, Mahaffy PR, Grebowsky JM, Fox JL, Yelle RV, Jakosky BM (2015) First measurements of composition and dynamics of the martian ionosphere by maven's neutral gas and ion mass spectrometer. Geophys Res Lett 42(21):8958–8965. doi:10.1002/2015GL066146

Bilitza D, Altadill D, Zhang Y, Mertens C, Truhlik V, Richards P, McKinnell L, Reinisch B (2014) The international reference ionosphere 2012 a model of international collaboration. J Space Weather Space Clim 4:A07. doi:10.1051/swsc/2014004

Chapman S (1931) The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth part ii. grazing incidence. Proc Phys Soc 43(5):483. http://stacks.iop.org/0959-5309/43/i=5/a=302

Chen RH, Cravens TE, Nagy AF (1978) The Martian ionosphere in light of the Viking observations. J Geophys Res 83(A8):647–664

Eleman F (1973) The geomagnetic field. In: Cosmical geophysics, pp 45–62
Elias AG, Zossi BS, Yiğ E, Saavedra Z, de Haro Barbas BF (2017) Earth's magnetic field effect on MUF calculation and consequences for hmf2 trend estimates. JASTP. http://dx.doi.org/10.1016/j.jastp.2017.03.004, http://www.sciencedirect.com/science/article/pii/S1364682616304229
Finlay C, Maus S, Beggan C, Bondar T, Chambodut A, Chernova T, Chulliat A, Golovkov V, Hamilton B, Hamoudi M et al (2010) International geomagnetic reference field: the eleventh generation. Geophys J Int 183(3):1216–1230
Finlay CC, Olsen N, Tøffner-Clausen L (2015) Dtu candidate field models for igrf-12 and the chaos-5 geomagnetic field model. Earth, Planets and Space 67(1):114. doi:10.1186/s40623-015-0274-3
Forbes JM, Roble RG, Fesen CG (1993) Acceleration, heating, and compositional mixing of the thermosphere due to upward propagating tides. J Geophys Res Space Physics 98(A1):311–321. doi:10.1029/92JA00442
Garcia RR, Solomon S (1985) The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere. J Geophys Res 90:3850–3868, implementation of Lindzen's parameterization into a two-dimensional dynamical model to study the effects of GWs in the MLT
Gjerloev J (2012) The supermag data processing technique. J Geophys Res Space Physics 117(A9)
Gong Y, Zhou Q, Zhang SD, Aponte N, Sulzer M, Gonzalez S (2013) The F region and topside ionosphere response to a strong geomagnetic storm at Arecibo. J Geophys Res 118:51775183. doi:10.1002/jgra.50502
Laštovička J (2006) Forcing of the ionosphere by waves from below. J Atmos Sol-Terr Phys 68:479–497
Laštovička J (2009) Lower ionosphere response to external forcing: a brief review. Adv Space Res 43:1–14
Liu AZ, Chester CS (2004) Vertical dynamical transport of mesopheric constituents by dissipating gravity waves. J Atmos Sol-Terr Phys 66:267–275. doi:10.1016/j.jastp.2003.11.002
Maus SS, Macmillan T, Chernova T, Choi S, Dater D, Golokov V, Lesur V, Lowes F, Lühr H, Mai W, McLean S, Olsen N, Rother M, Sabaka T, Thomson A, Zvera T (2005) The 10th generation international geomagnetic reference field. Phys Earth Planetary Interiors 151:320–322
Nair H, Allen M, Anbar AD, Yung YL, Clancy RT (1994) A photochemical model of the martian atmosphere. Icarus 111:124–150
Offermann D, Jarischa M, Schmidt H, Oberheide J, Grossmann KU, Guseva O, Russell JM, Mlynczak MG (2007) The "wave turbopause". J Atmos Sol-Terr Phys 69:2139–2158. doi:10.1016/j.jastp.2007.05.012
Olsen N, Hulot G, Lesur V, Finlay CC, Beggan C, Chulliat A, Sabaka TJ, Floberghagen R, Friis-Christensen E, Haagmans R, Kotsiaros S, Lhr H, Tffner-Clausen L, Vigneron P (2015) The swarm initial field model for the 2014 geomagnetic field. Geophys Res Lett 42(4):1092–1098. doi:10.1002/2014GL062659
Pancheva D, Miyoshi Y, Mukhtarov P, Jin H, Shinagawa H, Fujiwara H (2012) Global response of the ionosphere to atmospheric tides forced from below: Comparison between cosmic measurements and simulations by atmosphere-ionosphere coupled model gaia. J Geophys Res Space Phys 117(A7): doi:10.1029/2011JA017452
Ratcliffe JA (1972) An introduction to the ionosphere and magnetosphere. Cambridge University Press, Cambridge
Ratcliffe JA, Weekes K (1960) The ionosphere. In: Ratcliffe JA (ed) Physics of the upper atmosphere, Academic press, pp 378–456
Reinisch B (2000) Radio sounding of geospace plasmas. Física de la Tierra 12:105
Rishbeth H, Garriott OK (1969) Introduction to ionospheric physics, International geophysics series, vol 14. Academic Press
Rishbeth H, Mendillo M (2004) Ionospheric layers of mars and earth. Planet Space Sci 52. doi:10.1016/j.pss.2004.02.007

Schunk RW, Nagy AF (2009) Ionospheres: Physics, plasma physics and chemistry. Atmospheric and space science series. Cambridge University Press

Singer SF, Maple E, Bowen WA (1951) Evidence for ionosphere currents from rocket experiments near the geomagnetic equator. J Geophys Res 56(2):265–281. doi:10.1029/JZ056i002p00265

Thébault E, Finlay CC, Beggan CD, Alken P, Aubert J, Barrois O, Bertrand F, Bondar T, Boness A, Brocco L, Canet E, Chambodut A, Chulliat A, Coïsson P, Civet F, Du A, Fournier A, Fratter I, Gillet N, Hamilton B, Hamoudi M, Hulot G, Jager T, Korte M, Kuang W, Lalanne X, Langlais B, Léger JM, Lesur V, Lowes FJ, Macmillan S, Mandea M, Manoj C, Maus S, Olsen N, Petrov V, Ridley V, Rother M, Sabaka TJ, Saturnino D, Schachtschneider R, Sirol O, Tangborn A, Thomson A, Tøffner-Clausen L, Vigneron P, Wardinski I, Zvereva T (2015) International geomagnetic reference field: the 12th generation. Earth Planets Space 67(1):79. doi:10.1186/s40623-015-0228-9

Vichare G, Ridley A, Yiğit E (2012) Quiet-time low latitude ionospheric electrodynamics in the non-hydrostatic global ionospherethermosphere model. J Atmos Sol-Terr Phys 80:161–172. doi:10.1016/j.jastp.2012.01.009

Walterscheid RL, Hickey MP (2012) Gravity wave propagation in a diffusively separated gas: effects on the total gas. J Geophys Res 117:A05303. doi:10.1029/2011JA017451

Witasse O, Cravens T, Mendillo M, Moses J, Kliore A, Nagy A, Breus T (2008) Solar system ionospheres. Space Sci Rev 139:235–265. doi:10.1007/s11214-008-9395-3

Withers P, Vogt M, Mahaffy P, Benna M, Elrod M, Jakosky B (2015) Changes in the thermosphere and ionosphere of mars from viking to maven. Geophys Res Lett 42. doi:10.1002/2015GRL065985

Yiğit E (2015) Atmospheric and space sciences: neutral atmospheres, vol 1. SpringerBriefs in Earth Sci., Springer, Netherlands. doi:10.1007/978-3-319-21581-5

Yiğit E, Ridley AJ (2011) Role of variability in determining the vertical wind speeds and structure. J Geophys Res 116:A12305. doi:10.1029/2011JA016714

Yiğit E, Ridley AJ, Moldwin MB (2012) Importance of capturing heliospheric variability for studies of thermospheric vertical winds. J Geophys Res 117:A07306. doi:10.1029/2012JA017596

Yiğit E, Knížová PK, Georgieva K, Ward W (2016) A review of vertical coupling in the atmosphere-ionosphere system: effects of waves, sudden stratospheric warmings, space weather, and of solar activity. J Atmos Sol-Terr Phys 141:1–12. doi:10.1016/j.jastp.2016.02.011, http://www.sciencedirect.com/science/article/pii/S1364682616300426

Yonezawa T (1966) Theory of formation of the ionosphere. Space Sci Rev 5:3–56

Chapter 5
Dynamics of the Atmosphere-Ionosphere System

Meteorological Influences, Variability, and Space Weather

The fundamental idea is that atmospheric pressures, velocities, etc. should be expressed as numbers, and should be tabulated at certain latitudes, longitudes and heights, so as to give a general account of the state of the atmosphere at any instant, over an extended region, up to a height of say 20 km. The numbers in this table are supposed to be given, at a certain initial instant, by means of observations.
It is shown that there is an arithmetical method of operating upon these tabulated numbers, so as to obtain a new table representing approximately the subsequent state of the atmosphere after a brief interval of time, δt say. The process can be repeated so as to yield the state of the atmosphere after $2\delta t$, $3\delta t$ and so on. There is a limit however to the possible number of repetitions, because each table is found to be smaller than its predecessor, in longitude and latitude, having lost a strip round its edge. Only if the table included the whole globe could the repetitions be endless. Also the errors increase with the number of steps.

Richardson in "Weather Prediction by Numerical Process" (1922).

Abstract The upper atmosphere is the outermost region of a planetary atmosphere where solar radiation leads to formation of an ionosphere with appreciable plasma densities. The upper atmosphere and ionosphere continuously interact with each other in a two-way nonlinear manner. It is therefore increasingly common to refer to them as the atmosphere-ionosphere system. Overall, the state and evolution of the atmotsphere-ionosphere system is influenced by meteorological processes from below, which is called "coupling from below", and by solar and geomagnetic processes, often defined as "coupling from above" or simply space weather. Meteorological effects include upward propagating gravity waves, planetary waves and tides from the lower atmosphere into the thermosphere. This chapter reviews some elements of energetics and dynamics of the thermosphere-ionosphere system. Gravity waves are shown to play an appreciable role for the energy and momentum budget of the thermosphere, while planetary waves modulate lower thermospheric variability over long time scales. During geomagnetic storms, which are a prominent manifestation of space weather, the thermosphere and ionosphere undergo substantial

E. Yiğit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-62006-0_5

dynamical and thermal changes due to enhanced ion drag, Joule heating, and particle precipitation. A comprehensive understanding of the response of the atmosphere-ionosphere system to processes of lower atmospheric origin and those originating at Sun require a whole atmosphere approach.

Keywords Atmosphere-ionosphere dynamics · Variability · Vertical coupling · Gravity waves · Tides · General circulation modeling · Thermosphere-ionosphere · Space weather · Geomagnetic storms

5.1 Introduction

The atmosphere-ionosphere system is influenced by a combination of dynamical, energetic, and mixing processes. It is influenced by meteorological phenomena from below (Laštovička 2006; Yiğit et al. 2009, 2012a, 2014, 2016b; Yiğit and Medvedev 2015, 2016; Hickey et al. 2009, 2010; Kazimirovsky et al. 2003; Galvan et al. 2012; Vadas et al. 2009; Oberheide et al. 2015) and by solar and geomagnetic processes and solar variability (Yiğit et al. 2016a; Huang et al. 2014; Blanch et al. 2013; Gong et al. 2013; Balan et al. 2011; Mikhailov 2008) from above as illustrated in Fig. 5.1 in a simplified fashion. Meteorological processes include the direct effects of internal waves propagating from below (gravity waves, tides, planetary waves, etc.) and other transient events, such as sudden stratospheric warmings (SSWs) (Goncharenko et al. 2013; Fagundes et al. 2015; Miyoshi et al. 2015; Yiğit and Medvedev 2012) and quasi-biennial oscillations (QBOs) (Canziani and Holton 1998; Holton and Lindzen 1972). It is also common to relate the variations of the thermospheric and/or ionospheric parameters to specific lower atmospheric processes, such as earthquakes, convection, thunderstorms, tsunamis, etc.

Coupling from above entails the ionizing effects of Sun, solar wind and coronal mass ejection (CMEs) effects, precipitating particles (Clilverd et al. 2015; Wilson et al. 2006; Baker et al. 2004; Galand et al. 2001), and electric fields of magnetospheric origin (Klimenko and Klimenko 2012; Deng et al. 2009; Matsuo and Richmond 2008). Solar wind propagating from Sun and impacting the terrestrial magnetosphere can cause severe geomagnetic storms in the ionosphere, which ultimately influence the upper atmosphere and other layers below. Severe magnetic storms can impact the surface and technological facilities.

The overall goal of this final chapter is to discuss some basic physical mechanisms and selected research topics of the atmosphere-ionosphere system. Previous chapters have so far reviewed various physical concepts of space physics and ionospheric chemistry. There are many interesting challenges in ionospheric physics. Especially, in a planetary science context, due to a growing interest in the aeronomy of other planets, ionospheric physics will concern many researchers who are interested in better understanding other planets. In particular, given the rapid technological advancements in atmospheric remote-sensing and in the development of complex global scale numerical models, science community is occupied with an ever increasing

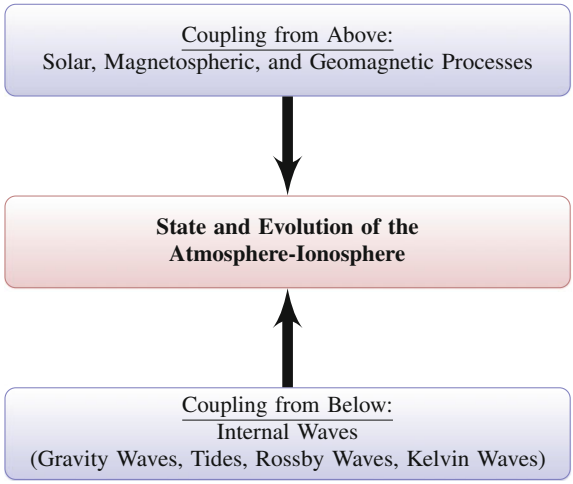

Fig. 5.1 A simple illustration of Vertical coupling in the atmosphere-ionosphere system. Coupling from above includes processes such as solar, magnetospheric, and geomagnetic origin, while coupling from below denotes the combined effects of meteorological processes. Adopted from (Yiğit et al. 2016b, Fig. 1)

number of fundamental science questions concerning the variations of atmospheric and ionospheric environments on different spatiotemporal scales.

On the one hand, satellite observations provide a detailed survey of the upper atmosphere, where they can operate in a highly rarefied atmosphere. On the other hand, complex atmospheric models called general circulation models (GCMs) (Chap. 6 Yiğit 2015) that solve in an ever increasing spatial resolution the fundamental conservation equations for every minute of the day for all locations of the globe are being developed. It has become a common research activity to use models that have been extensively developed and tested in Earth's atmosphere for other planetary atmospheres and ionospheres. I think, it is therefore instructive to start with the discussion of modeling of the dominant physical mechanisms in the upper atmosphere, which the next section will focus on. In this context, it is important to realize that the ultimate goal of atmospheric modeling is not to account for every possible process that is occurring in nature. It is in fact pretty impossible to achieve this goal. Some small-scale processes like gravity waves, convection, and clouds can often not be self-consistently represented in GCMs; hence they are parameterized. Improving physical foundations of such parameterizations may lead to an improved representation of nature. What however is a key requirement for a successful modeling is to identify the important processes in the atmosphere and then formulate a mathematically accurate representation of these physical processes, which can then be converted to a robust computer model utilizing appropriate numerical techniques. One often analyzes the time scales over which physical processes occur in order to evaluate their relative significance. Often chemical and dynamical time constants are compared to each other. So, overall, modeling is a fascinating activity during which solutions of complex physical and chemical equations of physics are approximated via numerical methods for the purpose of better understanding how nature works. In summary, the following three science questions, among many others, can help

appreciate the complexity of the planetary atmospheres and ionospheres and guide one through this challenging endeavor:

- What are the important key physical processes (internal and external) in the atmosphere-ionosphere system?
- How do these physical processes control the state and evolution of the system?
- How are the different layers within this system are coupled to each other?

Next section presents some basic principles of modeling planetary atmospheres and ionospheres. Later sections will discuss meteorological influences on the upper atmosphere (Sect. 5.3), thermospheric vertical winds (Sect. 5.4), space weather and geomagnetic storms (Sects. 5.5 and 5.6).

5.2 Modeling Planetary Upper Atmospheres and Ionospheres

The purpose of modeling in the context of planetary atmospheres and ionospheres is to predict the future state of atmospheric dynamics from knowledge of its present state by using numerical methods in order to solve the coupled conservation equations. The feasibility of such an approach has already been proposed by Richardson (1922). The associated equations form a closed set of prediction equations, called prognostic equations, which relate field variables to each other. Next sections provide some insight into the dynamics and energetics of planetary atmospheres in the context of GCMs.

5.2.1 General Circulation Modeling of the Atmosphere-Ionosphere System

Models are powerful numerical tools in which various physical and chemical processes can be investigated in a controlled manner and their importance for a physical system can be assessed. One of the most sophisticated flavors of numerical models that are developed for the atmosphere-ionosphere system is GCMs. It is used to predict the future state of the atmospheric circulation from knowledge of its present state by using *numerical approximation techniques* to the dynamical and chemical equations based on conservation laws of physics. A comprehensive GCM uses the principles of energy, momentum, and mass conservation, taking into account various radiative, chemical, and dynamical processes, in order to simulate the state and evolution of the atmosphere. While doing so, it can be used to diagnose physical processes influencing the system. The assumed equations constitute a closed set of equations. These coupled equations are solved in a global grid, consisting of a discrete number of latitudes and longitudes in the horizontal direction, at fixed vertical

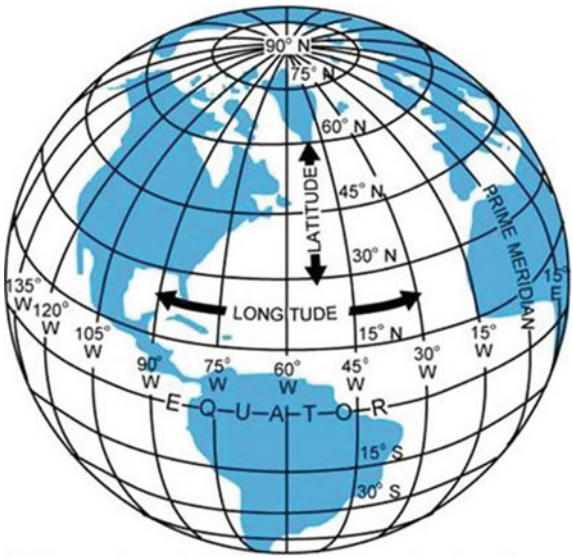

Fig. 5.2 An illustration of a representative horizontal grid used in general circulation models (GCMs), consisting of a discrete number of latitude (North-South direction) and longitude (East-West direction) grid points based on a specified resolution. The equator (0° latitude) and the Prime Meridian (0° longitude)

levels starting either at the planetary surface or at a reference level in the atmosphere. This three-dimensional space is called the *model domain*. An illustration of a horizontal numerical grid is seen in Fig. 5.2 where the Equator and the Prime Meridian are marked as representative of 0° latitude and 0° longitude, respectively. The longitude is aligned in the East-West direction (positive east), while the latitude is in the North-South direction (positive north). The horizontal model grid can then be represented by $m \times n$ latitude-longitude grid points. The higher the resolution of the model, that is, the larger the $m \times n$ grid points, the larger the horizontal resolution of the model and higher is the computational cost.

In a GCM, differential equations of physics and chemistry are then converted to algebraic difference equations. For example, for a given time step, the zonal inertial acceleration $u\,(\partial u/\partial x)$ can be approximated with a first-order forward difference equation as

$$u\frac{\partial u}{\partial x} \approx u(x, y, t)\left[\frac{u(x+\delta x, y, t) - u(x, y, t)}{\delta x}\right], \tag{5.1}$$

where δx is the grid spacing (i.e., resolution) in the x-direction. From the initial state $t = 0$, for which the initial conditions are specified, the dynamical equations are integrated in time to get future state $t > 0$, assuming specific boundary conditions. Coupled equations are solved for the particular grid points associated with a resolution. All atmospheric models require specification of boundary conditions in the lower boundary (e.g., the planetary surface) and in the upper boundary (i.e., at the top of the model domain). In practical applications, boundary conditions are chosen depending on the available observational constraints.

In the vertical direction, GCMs can assume a number of different vertical coordinate systems, each of which has various advantages and disadvantages. The specific choice of a coordinate system depends on the particular part of the atmosphere to be modeled. If the lower atmosphere is to be modeled in detail, then it is desirable that the vertical coordinates are terrain following, such as the σ-coordinate system. More sophisticated models extending from the surface to the middle and upper atmosphere utilize a hydrid-coordinate system, which is thought to combine the strengths of more than one vertical coordinate system. For example, the σ-p-coordinate (sigma-pressure coordinate) system (Simmons and Burridge 1981) is terrain-following in the lower atmosphere and it gradually transitions to a pressure coordinate system at higher levels. Various numerical weather forecast models and ensemble average models, such as the Goddard Earth Observing System-5 (GEOS-5) atmospheric general circulation model (AGCM) (Rienecker and CoAUTHORs 2008) and NCEP (National Centers for Environmental Prediction) model use this technique. Also, in the Max Planck Institute Martian General Circulation Model ((MPI-MGCM), Medvedev et al. 2013, 2015, 2016) and the Kühlungsborn Mechanistic Circulation Model ((KMCM), Zülicke and Becker 2013; Becker et al. 2015) this approach is implemented.

An example of model analysis at fixed pressure levels is shown in Fig. 5.3. Pressure level-latitude distribution of geopotential height Z is plotted for low solar activity conditions using the Coupled Middle Atmosphere Thermosphere-2 Earth GCM

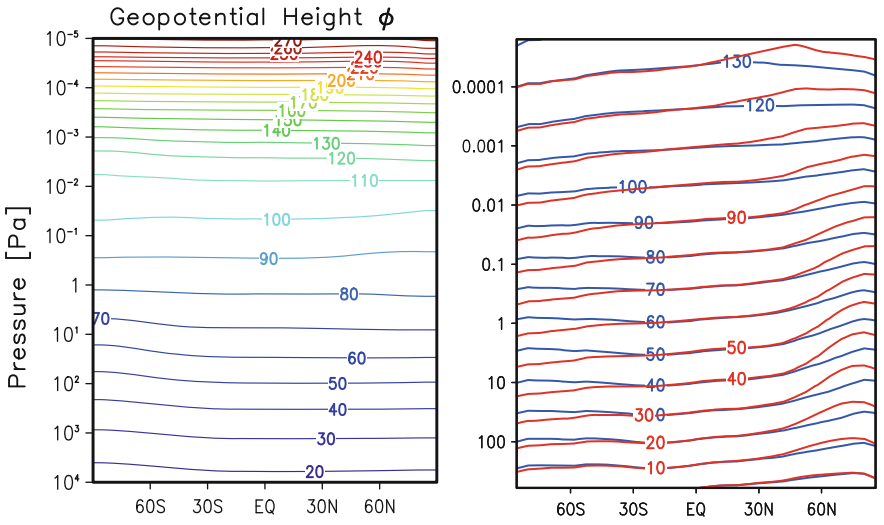

Fig. 5.3 Zonal mean geopotential height (Z) variations [km] as function of pressure [Pa] and latitude simulated by the Coupled Middle Atmosphere Thermosphere-2 Earth GCM (*left panel*) and by the Max Planck Institute for Solar System Research Martian GCM (Medvedev et al. 2011, Fig. 1). In the case of Mars, simulations are shown for $L_s = 180°$ (*blue*) and $L_s = 270°$ (*red*)

(left panel) and the MPI-MGCM (right panel). Note that the CMAT2 extends from 15 km (10^4 Pa) upward while the lower boundary of the MGCM is at the surface (~600 Pa).

CMAT2 utilizes the pressure coordinate system, where the vertical levels are specified as a function of scale height

$$p = p_0 \exp(-z/H) = p_0 \, \mathrm{d}n(1 - n), \tag{5.2}$$

where $\mathrm{d}n = 1/3$ states that each pressure level is separated from the next level by one-third of a scale height. The geopotential height is in general varies with season, solar and geomagnetic activity. If we compare these distributions to each other in order to find out to what pressure level Z corresponds to, we see that Mars' atmosphere is overall less dense at a given height as $p \propto \rho$. The pressure conditions on the surface of Mars are found in the stratosphere-lower mesosphere region of Earth at an altitude of 50–60 km.

It has been approximately half a century that GCMs have been used to characterize the large-scale structure of the thermosphere and ionosphere. Such models tried to isolate the thermosphere by assuming lower boundaries somewhere around the mesopause region (~90) km. One of the earlier models that resembled a simplified GCM was the two-dimensional steady-state thermospheric dynamics model of Dickinson et al. (1971). Later, more complex three-dimensional time-dependent GCMs were developed (Dickinson et al. 1981; Fuller-Rowell 1981). Further studies have realized the importance of self-consistent coupling between the ionosphere and thermosphere in order to better simulate the upper atmosphere. The Global Self-consistent Model of the Thermosphere, Ionosphere, Protonosphere (GSM-TIP) (Namgaladze et al. 1988) allows modeling studies accounting for various neutral and ionospheric drivers (Klimenko and Klimenko 2012).

Overall, modeling physical processes involves various simplifications, simultaneously illustrating how complex physics and chemistry can be approximated for the purpose of better understanding the large-scale structure of the atmosphere-ionosphere system.

Next I would like to discuss the complexity and fundamental physical processes that form the foundations of such models.

5.2.2 Complexity

As discussed in previous chapters, the atmosphere-ionosphere system is a complex multi-species environment. There are multiple neutral and ionized chemical species. The underlying plasma and neutral equations are highly nonlinear and can thus not be solved analytically. Only very simple estimations can be made, which can help better predict the future state of the system. Observations provide a method of characterizing the system but detailed understanding of the controlling physical

mechanisms require some modeling efforts. So, what do we mean by modeling? And what kind of physical insight can the modeling technique provide for the scientist?

As the first volume of my book focused on the modeling of the neutral atmosphere, or the general circulation modeling of the neutral atmosphere, I would like to extend the discussion of modeling to the ionosphere-thermosphere system. Essentially, the major component of the upper atmosphere, above the mesopause, is the thermosphere, the hottest region of a planetary atmosphere. which is much denser than the ionosphere in most planets studied so far but the ionosphere has a substantial dynamical and thermal influence on the neutral atmosphere via collisional energy and momentum transfer processes and vice versa. This type of coupling is referred to as a two-way coupling, which is a highly nonlinear process. GCMs that are designed to simulate the atmosphere-ionosphere system are based on the fluid model of the atmospheric gas, as discussed in Sect. 3.1.

Modeling the ionosphere can be structured by considering various submodels or components, such as, the neutrals, ion production and loss, thus ion chemistry, transport, electric fields and magnetic field models. These components are the key components of a theoretical ionosphere model (Nisbet 1975). Here I would like to highlight some of the selected key physical mechanisms that are particularly important for the energy and momentum budget of the thermosphere-ionosphere system by organizing the discussion as aspects of dynamics and energetics.

5.2.3 *Dynamics*

In the context of dynamics one is primarily interested in the momentum balance, which can be expressed by Newton's law:

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \sum_i \mathscr{F}_i, \tag{5.3}$$

stating that the net acceleration acting on a body is given by the sum of the various body forces and viscous forces per unit mass $\mathscr{F}_i$ acting on the body. In a dynamical equilibrium state the sum of all forces acting on the system vanishes

$$\sum_i \mathscr{F}_i = 0, \tag{5.4}$$

The more precisely the spatiotemporal variations of the contributing forces are represented, the more accurate is the modeling of thermospheric and ionospheric dynamics. Note that the governing equations have to be written both for the neutral atmosphere and the ionosphere as the upper atmosphere is a two-component fluid—neutrals and charged particles (ions and electrons).

On the one hand, presence of an ionosphere modifies the general circulation of the thermosphere, and on the other hand the neutral atmosphere impacts the ionosphere significantly. In the ionosphere, the ion concentration can reach significant levels so that the ion motion can impact the neutral air motion. This frictional effect of the ions on the neutral is called the *ion drag* and is a major physical mechanism, in particular, in the F region. Given that ν_{ni} expresses the neutral-ion collision frequency, that is, the frequency of the collisions of the neutrals with the surrounding ions, the body force per unit mass $\mathbf{a}_i$ acting on the neutrals due to the neutral-ion collisions is given by

$$\mathbf{a}_i = \left(\frac{\partial \mathbf{u}_n}{\partial t}\right)_{ni} = -\nu_{ni}(\mathbf{u}_n - \mathbf{u}_i). \tag{5.5}$$

We have the relationship from 2.62

$$\rho_i \nu_{in} = \rho_n \nu_{ni} \tag{5.6}$$

to express the ion drag in terms of the densities as

$$\mathbf{a}_i = -\nu_{in}\frac{\rho_i}{\rho_n}(\mathbf{u}_n - \mathbf{u}_i), \tag{5.7}$$

where in a multi-component planetary atmosphere and ionosphere, chemical species are diffusively separated in the vertical direction and we have different densities for the different species and thus for the total ion and neutral mass densities we get

$$\rho_i = \sum_{i_s} \rho_{i_s} = \sum_{i_s} n_{i_s} m_{i_s} \tag{5.8}$$

$$\rho_n = \sum_{n_s} \rho_{n_s} = \sum_{n_s} n_{n_s} m_{n_s}. \tag{5.9}$$

Above the subscripts "i_s" and "n_s" denote the different ion and neutral species in a planetary upper atmosphere. From the above relations, we can conclude that as the neutral atmosphere is much denser than the ionosphere $\rho_n > \rho_i$, we have $\nu_{in} > \nu_{ni}$, that is the collision frequency of the ions with the neutrals is larger than the collision frequency of neutrals with the ions. Then, are the effects of the neutrals on the ions the same as the effects of the ions on the neutrals? Not at all. If we write down the neutral drag term on the ions similar to the expression (5.5), we see that, because $\rho_n/\rho_i > 1$, the neutral drag is larger than and oppositely directed to the ion drag for a given configuration of relative ion-neutral (or neutral-ion) differential velocity.

It is seen that the neutral air is accelerated, if we assumed the ion velocity to be constant, with a time constant of $\tau_i = \nu_{ni}^{-1}$ toward the ion velocity.

Based on our discussion of the governing equations of atmospheric physics in the first volume of this monograph series and discussion of transport equations, we can state the complete momentum equation for a (neutral) species as

$$\frac{\mathrm{d}\mathbf{u}_s}{\mathrm{d}t} = -\frac{1}{\rho_s}\nabla p_s + \nu\,\triangle\mathbf{u}_s - 2\boldsymbol{\Omega}\times\mathbf{u}_s + \mathbf{g} - \sum_j \nu_{sj}(\mathbf{u}_s - \mathbf{u}_j) + \mathscr{F}'_s \quad (5.10)$$

where we have on the right hand side the pressure force, viscosity, Coriolis force, gravity, collisional drag, and small-scale gravity wave (GW) effects ($\mathscr{F}'$).

Specifically, we can write Eq. (5.10) for the thermosphere, noting that the different neutral species perform a drift motion together

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = -\frac{1}{\rho}\nabla p + \nu\,\triangle\mathbf{u} - 2\boldsymbol{\Omega}\times\mathbf{u} + \mathbf{g} - \nu_{ni}(\mathbf{u} - \mathbf{u}_i) + \mathscr{F}', \quad (5.11)$$

where we have used the neutral-ion collision frequency ν_{ni}, neutral drift velocity $\mathbf{u}$, and ion drift velocity $\mathbf{u}_i$. The collisional drag for neutrals is caused mainly by ions and is thus called ion drag. Various approximations are possible to the momentum equation, such as the geostrophic and gradient wind balances. The corresponding momentum equations for electrons and ions are as follows:

$$\frac{\mathrm{d}\mathbf{u}_i}{\mathrm{d}t} = \frac{1}{\rho_i}\nabla\cdot\sigma_i - 2\boldsymbol{\Omega}\times\mathbf{u}_i + \mathbf{g} + \frac{q_i}{m_i}(\mathbf{E} + \mathbf{u}_i\times\mathbf{B}) - \sum_s \nu_{is}(\mathbf{u}_i - \mathbf{u}_s), \quad (5.12)$$

$$\frac{\mathrm{d}\mathbf{u}_e}{\mathrm{d}t} = \frac{1}{\rho_e}\nabla\cdot\sigma_e - 2\boldsymbol{\Omega}\times\mathbf{u}_e + \mathbf{g} + \frac{q_e}{m_e}(\mathbf{E} + \mathbf{u}_e\times\mathbf{B}) - \sum_s \nu_{es}(\mathbf{u}_e - \mathbf{u}_s), \quad (5.13)$$

where σ_i and σ_e denote the ion and electron stress tensor terms, respectively, as defined in a generic form in Sect. 3.9, and the fourth term on the right-hand side is the Lorentz force term, and the final term is the collision term that accounts for the interactions of the electrons and ions with the other species. In the *steady-state approximation*, the Lorentz force balances the collisional interactions, assuming that the major collisional momentum transfers occur due to the interactions with the ambient neutrals:

$$-\frac{e}{m_i}(\mathbf{E} + \mathbf{u}_i\times\mathbf{B}) = -\nu_{in}(\mathbf{u}_i - \mathbf{u}_n), \quad (5.14)$$

$$\frac{e}{m_e}(\mathbf{E} + \mathbf{u}_e\times\mathbf{B}) = -\nu_{en}(\mathbf{u}_e - \mathbf{u}_n), \quad (5.15)$$

where we made use of the elementary charge e for $q_i = e$ and $q_e = -e$.

Models can be readily used to analyze the contribution of different forces to the momentum balance of plasma and neutrals. For example, Kwak and Richmond (2007) have used a GCM to study the dynamics of thermospheric winds below 170 km for a negative IMF B_y conditions. Their results indicated that winds above 123 km were maintained by the gradient-wind balance among the pressure gradient

force, Coriolis, and horizontal momentum advection forces. Yiğit et al. (2009) have modeled that the small-scale gravity wave (GW) effect term $\mathscr{F}'$ as parameterized by the scheme of Yiğit et al. (2008) compete with the ion drag term up to F region altitudes in a zonal mean sense.

5.2.4 Energetics

Energy is conserved. Therefore, according to the first law of thermodynamics, the change of internal energy per unit mass inside a fluid element must be equal to the net gain of energy (heat flow into the system) plus the work done per unit mass on the fluid element due to body forces

$$\mathrm{d}U = \mathrm{d}Q - \mathrm{d}W, \tag{5.16}$$

where $\mathrm{d}Q > 0$ represents heat flow into the system, $\mathrm{d}W = p\mathrm{d}V$, and $\mathrm{d}W > 0$ is the work done by the system. Accordingly, work being performed by the system contributes toward reducing the internal energy of the system, while performing work on the system $\mathrm{d}W < 0$ increases its internal energy. Equation (5.16) can be formulated with the equation of state $p = \rho RT/M$ and entropy change $\mathrm{d}S = \mathrm{d}Q\ T$ to write energy equation in terms of pressure and temperature as

$$T\frac{\mathrm{d}S}{\mathrm{d}t} = c_p\frac{\mathrm{d}T}{\mathrm{d}t} - \frac{1}{\rho}\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\mathrm{d}Q}{\mathrm{d}t}, \tag{5.17}$$

where $\mathrm{d}Q/\mathrm{d}T$ is the net heating rate per unit mass (energy deposition rate in $\mathrm{W\,kg^{-1}}$) due to diabatic processes, which lead to a change of state due to energy exchange with the environments as a result of temperature difference.

For adiabatic processes, the system is thermally isolated from its environment, thus the entropy does not change:

$$T\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\mathrm{d}Q}{\mathrm{d}t} = 0, \tag{5.18}$$

and the energy balance is given by

$$c_p\frac{\mathrm{d}T}{\mathrm{d}t} = \frac{1}{\rho}\frac{\mathrm{d}p}{\mathrm{d}t}. \tag{5.19}$$

This relationship between the pressure and temperature expresses cooling via expansion or heating via compression associated with adiabatic processes, during which heat flow in or out of the system is neglected.

We shall discuss various processes, which contribute to the energy budget of the upper atmosphere. The mesopause with a mean temperature of around 120–130 K is the coldest spot in Earth's atmosphere as has been observed, for example, by falling sphere techniques (Lübken 1999) and modeled by middle atmosphere GCMs (Berger and von Zahn 1999). On the other hand, the thermosphere is the hottest region with temperatures ranging from ∼700 to ∼1500 K depending on solar and geomagnetic activity. How does the temperature in Earth's atmosphere increase so rapidly from the mesopause at around 90–100 km to several hundred K in the thermosphere? This feature is also present in Mars atmosphere but Mars' mesopause and thermosphere are colder than the terrestrial ones. The primary reason why the thermosphere is so hot is due to the combination of the absorption of solar UV/EUV energy by atmospheric constituents and the inefficient downward heat conduction. During the day, solar UV radiation is the most important heat source (Bates 1951). The production of heat by UV radiation depends primarily on the specific wavelength λ and the altitude in the atmosphere. Concentration of absorbing constituents n_s and the associated absorption cross sections σ_s, the solar energy flux at the top of the atmosphere Φ_{uv}, and the efficiency of the conversion of the UV energy to heat (i.e., heating efficiency) ε_s should be known. Then, a simple expression for the rate of heat production [W m^{-3}] due to UV absorption can be given by (Banks and Kockarts 1973)

$$q_{uv}(\lambda, z) = \sum_s \varepsilon_s(\lambda, z)\, \sigma_s(\lambda)\, n_s(z)\, \Phi_{uv}(\lambda)\, \exp[-\tau(\lambda, z)], \tag{5.20}$$

where τ is the wavelength and altitude dependent optical depth. I have discussed its concept in the ionospheric context in Sect. 4.7. The solar energy flux depends on the wavelength and flux of the photons outside the atmosphere. The absorption cross section [m^2] expresses the ability of a species to absorb a photon of wavelength λ. Overall, Eq. (5.20) expresses the energy deposition rate due to solar energy absorption by a given species "s".

In Earth's thermosphere, knowledge about the photoabsorption cross sections for O, N_2, and O_2 are most important. Earlier measurements of the solar EUV and UV fluxes in the Schumann-Runge continuum ($\lambda < 170$ nm) have been performed by the Atmosphere Explorer satellites in the 1970s. Besides radiative heating via photoabsorption, radiative cooling processes should be accounted for in order to accurately model thermospheric temperatures. The 63 μm emission by atomic oxygen (Bates 1951) and, especially, the 5.3 μm nitric oxide (NO) radiative emission (Kockarts 1980) and CO_2 radiative emission in the 15 μm (Fomichev et al. 1998) are important cooling mechanisms in Earth's thermosphere. In particular, the 5.3 μm NO cooling is temperature-dependent and has been modeled to be larger during solar maximum conditions than during minimum conditions (Roble and Emery 1983).

In addition to the radiative heating and cooling processes, the effects of Joule heating (Knipp et al. 2004), particle precipitation (Fuller-Rowell and Rees 1981;

Wilson et al. 2006), chemical heating (Kellogg 1961; Hickey et al. 2003), and small-scale GW effects (Medvedev and Klaassen 2003; Becker 2004; Yiğit and Medvedev 2009, 2013; Walterscheid et al. 2013) ought to be included in the energy budget of the thermosphere.

In a planetary atmosphere, the energy input varies strongly as a function of altitude. Considering these energetic processes, temperature variation can be estimated by the following approximation to the energy equation:

$$\rho c_p \left(\frac{\partial T}{\partial t} + w \frac{\partial T}{\partial z} \right) = \frac{\partial}{\partial z} \left(\kappa \frac{\partial T}{\partial z} \right) + Q - L_{IR} + \mathscr{E}_J + \mathscr{E}_p + \mathscr{E}', \tag{5.21}$$

where $c_p = R + c_v$ is specific heat at constant pressure, κ is temperature- and altitude-dependent thermal conductivity [J K^{-1} m^{-1} s^{-1}], $\mathscr{E}_J$ is Joule heating, $\mathscr{E}_p$ is heating due to particle precipitation, and $\mathscr{E}'$ represents the heating and/or cooling due to small-scale GWs in units of W m^{-3}.

5.3 Meteorological Influences on the Upper Atmosphere

In his conceptual article, Rishbeth (2002) has emphasized that not all ionospheric variability is caused by solar or geomagnetic processes and indicated that some variability may originate from below in form of internal waves. Additionally, he asked whether the lower atmospheric weather can affect the ionosphere. In general, weather substantially influences the state of the atmosphere. With the current state of knowledge gained, in particular, by global circulation models and space-bourne satellites, we can see that the question of whether the lower atmosphere influences the upper atmosphere or not is increasingly becoming a rhetorical question. The fact that the lower atmosphere is extremely variable and thus predictable only to a limited extent complicates this picture of vertical coupling.

Some weather phenomena are pretty common, for example, rain, snow, wind, fog, and clouds. However some are classified as natural weather phenomena that have a stronger potential to create lower atmospheric disturbances, for example, tornadoes, hurricanes, snow and wind storms. Figure 5.4 shows an image from the GOES-13 spacecraft observing the Hurricane Nicole on 12 October 2016 hitting Bahamas. GOES is a geostationary satellite at an altitude of 35,800 km and is able to obtain a full-disc view of Earth. This feature enables the monitoring of storm development and tracks. The majority of weather phenomena occurs in the troposphere. Overall, the main driver of the large-scale weather is the spatial difference in atmospheric pressure, temperature, and moisture.

Lower atmospheric variations produce a large spectrum of waves (Yiğit and Medvedev 2015). Small-scale gravity waves (GWs) of lower atmospheric origin propagate upward and influence the general circulation of the middle atmosphere (Holton 1982; Garcia and Solomon 1985; Garcia et al. 2007; Becker 2011) and thermosphere above the turbopause (e.g., Yiğit et al. 2009, 2012a; Miyoshi et al. 2014)

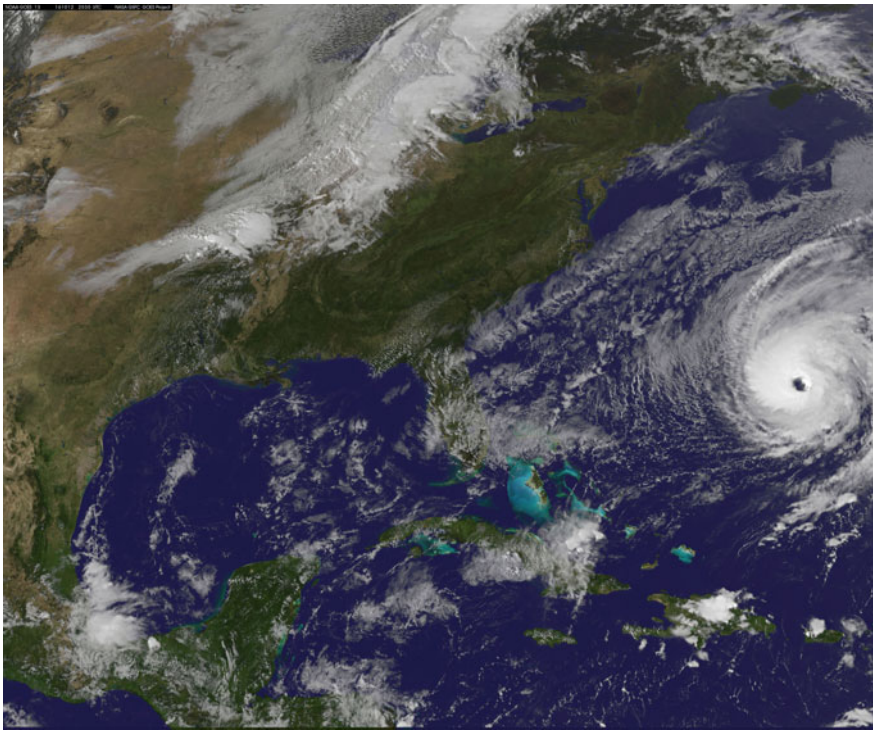

Fig. 5.4 Hurricane Nicole as observed on 12 October 2016 by NOAA's GOES-13 (Geostationary Operational Environmental Satellite-13) spacecraft. Credit: NASA-Goddard Space Flight Center, data from NOAA GOES

and modulate the ion-neutral coupling in the thermosphere-ionosphere (Yiğit et al. 2012a). In the thermosphere, the large-scale circulation, referred to as the general or global-scale circulation primarily consists of summer-to-winter flow of atmospheric air. This flow can be substantially modulated by ion-neutral friction force (i.e., ion drag, Eq. (5.5)) and heating of magnetospheric origin, such as particle precipitation and Joule heating (Roble et al. 1977; Thayer and Semester 2004).

The lower atmosphere is a source of a broad spectrum of internal waves. That is, from the perspective of the upper atmospheric physics, the lower atmosphere can be understood as the primary source of internal waves, in other words as the source of "primary waves". This property of the lower atmosphere means that a large number of waves of varying spatial scales (i.e., different wavelengths) are generated by various dynamical and meteorological processes, such as convection (Chun and Baik 2002; Beres et al. 2004). Earlier studies have indicated that internal gravity wave activity in the lower atmosphere can be adequately described by a universal spectrum. One example is seen in Fig. 5.5, where the power spectral density of mesoscale wind fluctuations as a function of frequency demonstrates a distinct power-law dependence (VanZandt 1982, Fig. 1).

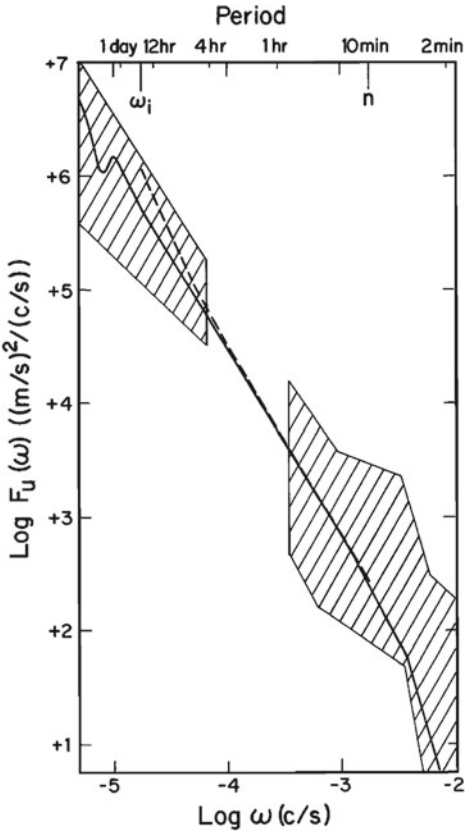

Fig. 5.5 Observed power spectral density of the horizontal winds (F_u) as a function of frequency ω shown by the *dashed curve*. The upper and lower limits of the hatched areas are the observed spectra used to evaluate the mean. The corresponding periods are shown in the upper x-axis ranging from a few days to minutes. (VanZandt 1982, Fig. 1)

The generated waves have the property to propagate upward, grow in amplitude upon penetrating the less denser upper atmospheric regions to conserve wave action, and ultimately encounter saturation at higher altitudes, where their wave saturation depends in a complex manner on the background atmosphere (stratification) and the individual wave parameters. Upward propagation of a large number of waves implies that the propagating waves have a distribution of phase speeds, momentum fluxes (or amplitudes), and scales. In principle, there are several thermal and dynamical wave generation mechanisms and each would produce a separate spectrum of waves that probably change in time. Most models assume a mean background GW spectrum that is representative of the composite effect of all generation mechanisms. In order to evaluate the effects of GWs, each wave in the spectrum should be treated individually and in regard to other waves, if nonlinear effects are taken into account. In fact, nonlinear effects produce additional damping on the waves (Weinstock 1982; Medvedev and Klaassen 2000; Yiğit et al. 2008). Essentially, upward propagating waves influence the background atmosphere upon gradual attenuation or dissipation, which describes the processes of wave saturation, and they continuously interact with each other.

The dynamics of internal waves can be summarized as the generation, propagation and interaction/saturation of waves. In order to adequately model the atmosphere on a global scale, physics and thermodynamics of internal waves have to be accurately represented. Thus, depending on the scales of internal waves, some waves are self-consistently resolved by models (e.g., Miyoshi et al. 2014; Kuroda et al. 2015), and some smaller scale waves have to be represented by parameterizations (e.g., Garcia et al. 2007; Yiğit et al. 2009). A wave parameterization should estimate both the sources and the effects (saturation) of these waves. Typically, parameterizations are simplified by assuming an empirical distribution of wave sources. Then, the biggest challenge remains to be the representation of the physics of wave propagation and saturation, i.e., interaction with the mean-flow.

Some well-known internal waves are solar tides, planetary waves, gravity waves, and Kelvin waves. Some of these waves possess various different "flavors", for example, in the case of solar tides, the migrating and nonmigrating tides or for GWs, small-scale and inertia-gravity waves are often studied. Solar (thermal) tides are primarily generated in the lower atmosphere by periodic absorption of solar radiation by tropospheric water vapor and stratospheric ozone. This absorption is strongly dependent on the local time (LT) as well as longitude and latitude.

As indicated in previous discussions, internal wave propagation from the lower atmosphere to upper atmosphere is the primary mechanism of how meteorological processes influence the thermosphere-ionosphere system. This coupling between the lower atmosphere and upper atmosphere is often called *vertical coupling*. This vertical coupling can occur in the upward or downward direction. That is, changes in the ionosphere can influence the atmosphere lower down as well. Recently, there is an increasing number of publications focusing on the vertical coupling in the atmosphere-ionosphere system (e.g., Pancheva et al. 2012; Knizova et al. 2015; Laskar et al. 2015; Oberheide et al. 2015; Tonev and Velinov 2015; Yiğit and Medvedev 2015; Yiğit et al. 2016b; Yiğit and Medvedev 2016). In fact, a number of previous studies have suggested that there may be an appreciable coupling between the lower and the upper atmosphere. The traditional ICMA/IAGA/SCOSTEP workshop on the Vertical Coupling in the Atmosphere-Ionosphere System has been regularly hosting presentations on the physical processes that couple the lower and upper atmosphere.

Direct propagation of GWs from the lower atmosphere into the thermosphere and the resulting momentum deposition in the upper atmosphere can be readily studied by GCMs. Figure 5.6 presents the mean zonal mean GW drag (upper panels) and ion drag (lower panels) in the "cut-off" (left column) and "extended" (right column) runs simulated by the CMAT2 model for September (equinox) conditions during low solar and geomagnetic activity. In the cut-off simulation GW propagation from the lower atmosphere to the upper atmosphere is not included beyond the turbopause ($\sim$105 km). In the extended simulation, lower atmospheric GWs are permitted to propagate into the upper thermosphere. Intercomparison of the cut-off and the extended simulations is a method of determining how much GWs propagate into the thermosphere and influence the circulation there. Essentially, if GWs do not propagate into the thermosphere, which has been the traditional assumption in GCMs, then

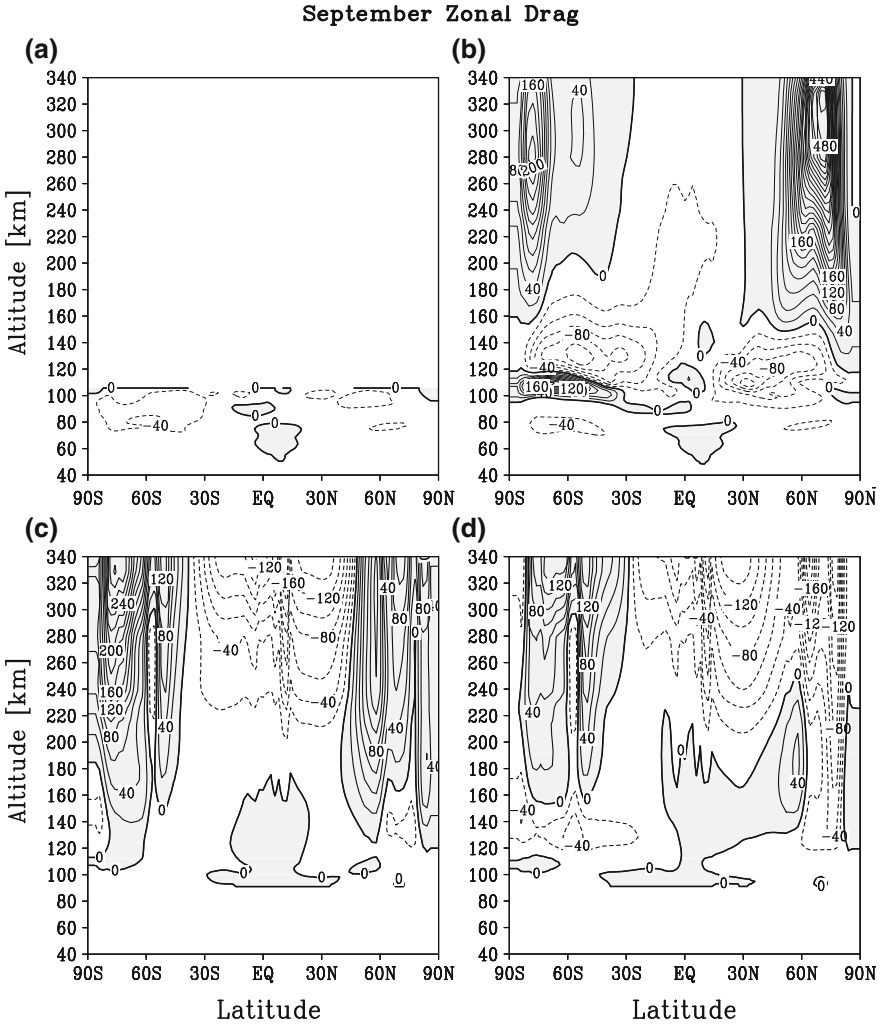

Fig. 5.6 Mean zonal mean gravity wave (GW) drag (*upper panels*) and ion drag (*lower panels*) in the cut-off (*left column*) and extended (*right column*) simulations performed by the CMAT2 model for September (equinox) conditions at low solar activity. Contour intervals are 20 m s^{-1} day^{-1}. Solid contour lines with gray shading represent eastward (positive) drag and dashed contour lines show the regions of westward (negative) drag. CMAT2 model uses the extended nonlinear spectral GW parameterization of Yiğit et al. (2008) to calculate GW drag and the Parameterized Ionosphere Model ((PIM), Daniell et al. 1995) to calculate ion drag. One month average is shown. (Yiğit et al. 2012a, Fig. 3)

we would not see any momentum flux divergence above the turbopause height. However, it is seen that not only GWs propagate into the thermosphere beyond the middle atmosphere and produce appreciable zonal body force there, they substantially influence the coupling between the ions and the neutrals (i.e., ion drag in Eq. (5.5)) by directly influencing the neutral drift motion by direct momentum deposition and thus modulating, in the shown case, the zonal component, $u_n - u_i$, of the relative neutral-ion drift velocity term, $(\mathbf{u}_n - \mathbf{u}_i)$, in the ion drag equation. Note that GW dynamical effects of up to $\pm$ 120 m s^{-1} day^{-1} in the lower thermosphere and 200 to 400 m s^{-1} day^{-1} in the high-latitude thermosphere are much stronger than the middle atmospheric effects of GWs, which is around 40 m s^{-1} day^{-1} in both hemispheres. Spectacularly, at high-latitudes in a zonal mean sense, thermospheric effects of GWs can locally exceed the mean ion drag effects, which are up to $\pm$100–200 ms^{-1} day^{-1}. These results are in agreement with the previous general circulation modeling studies of Yiğit et al. (2009) on the dynamical importance of lower atmospheric small-scale GWs for the general circulation of the thermosphere.

Planetary waves and solar tides significantly influence the morphology of the upper atmosphere. Figure 5.7 shows the altitude-time distribution of the amplitude of the ~5-day planetary wave with zonal wavenumber 1 as observed by the SABER

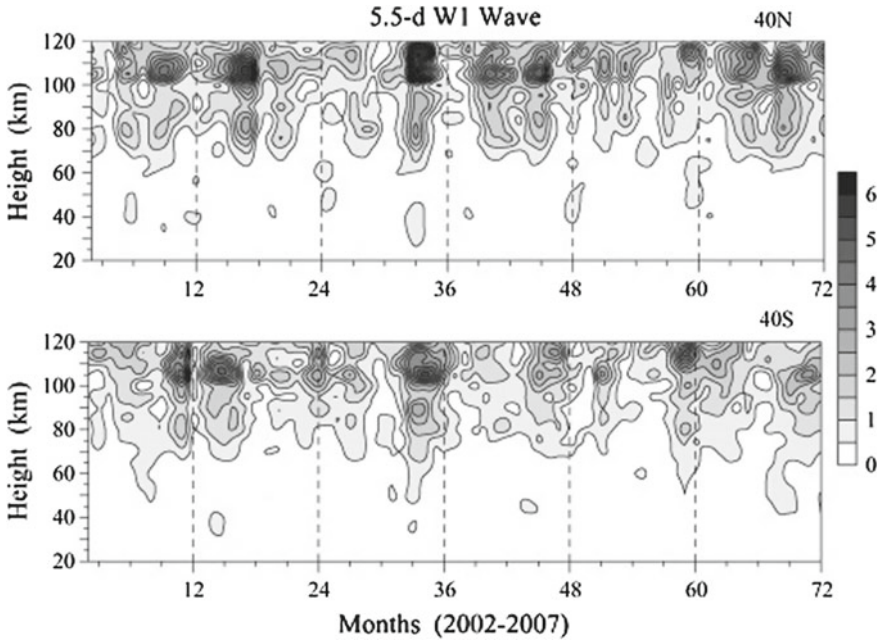

Fig. 5.7 Altitude-time cross sections of the 5.5-day Rossby wave amplitudes [K] at 40° N (*upper panel*) and 40° (*bottom panel*) as derived from SABER instrument on board the TIMED satellite from February 2002 to December 2007 (Fig. 3, Pancheva et al. 2010)

instrument on board TIMED satellite and analyzed in the work by Pancheva et al. (2010). They have studied data from February 2002 to December 2007. The wave amplitudes grow rapidly above 70 km, maximizing in the lower thermosphere with 5–6 K amplitudes during equinoctial periods. Observations have suggested that the diurnal tide is modulated by the planetary waves in the MLT (Guharay et al. 2015).

5.4 Thermospheric Vertical Winds and Coupling to the Ionosphere-Magnetosphere

The upper atmosphere demonstrates variability in various spatiotemporal scales. An exemplification of such variability is shown in Fig. 5.8, which shows GCM simulations of Yiğit et al. (2012b) of thermospheric vertical wind acceleration, vertical winds, auroral and Joule heating, and horizontal wind divergence from the upper panel downward. Results are shown for a two-hour period at a representative auroral latitude at 285 km during April. It is seen that the vertical wind demonstrates substantial temporal variability. Its hourly variation is influenced by the horizontal divergence, which indicates the importance of the large-scale circulation, and its much smaller minute-to-minute variations are influenced by the nonhydrostatic accelerations. Similar small-scale variations are seen in the heating terms as well. 5-min and 12-min smoothing results are with respect to the smoothing of the IMF input into the model, suggesting that variability in the IMF has a direct impact on the vertical winds.

5.5 Space Weather Effects

The space weather comprises the influences of Sun and geomagnetic processes on the geospace environment. More specifically, it is the effects of processes originating at Sun on Earth's atmospheric and magnetic field environment. The effects of the space weather can be severe. It can affect the biosphere, technological facilities, and day-to-day telecommunication systems. For example, essential communication and public activities depend on the Global Satellite Navigation System (GNNS), which is affected by scintillation in the ionosphere (Sreeja 2016). Space weather processes can impact the ionosphere and can thus severely influence GNNS. Disturbances of the geomagnetic field can cause geomagnetically induced currents in power grids and railway systems (Pulkkinen et al. 2017). Therefore, quantification of space weather effects and how we can predict their occurrences are important tasks for space physicists.

Essentially, our Sun affects the geospace environment via two major paths: radiation and energetic particles. Sun emits energetic plasma, which we refer to as the solar wind, into the heliosphere, which directly impacts the geospace. The solar

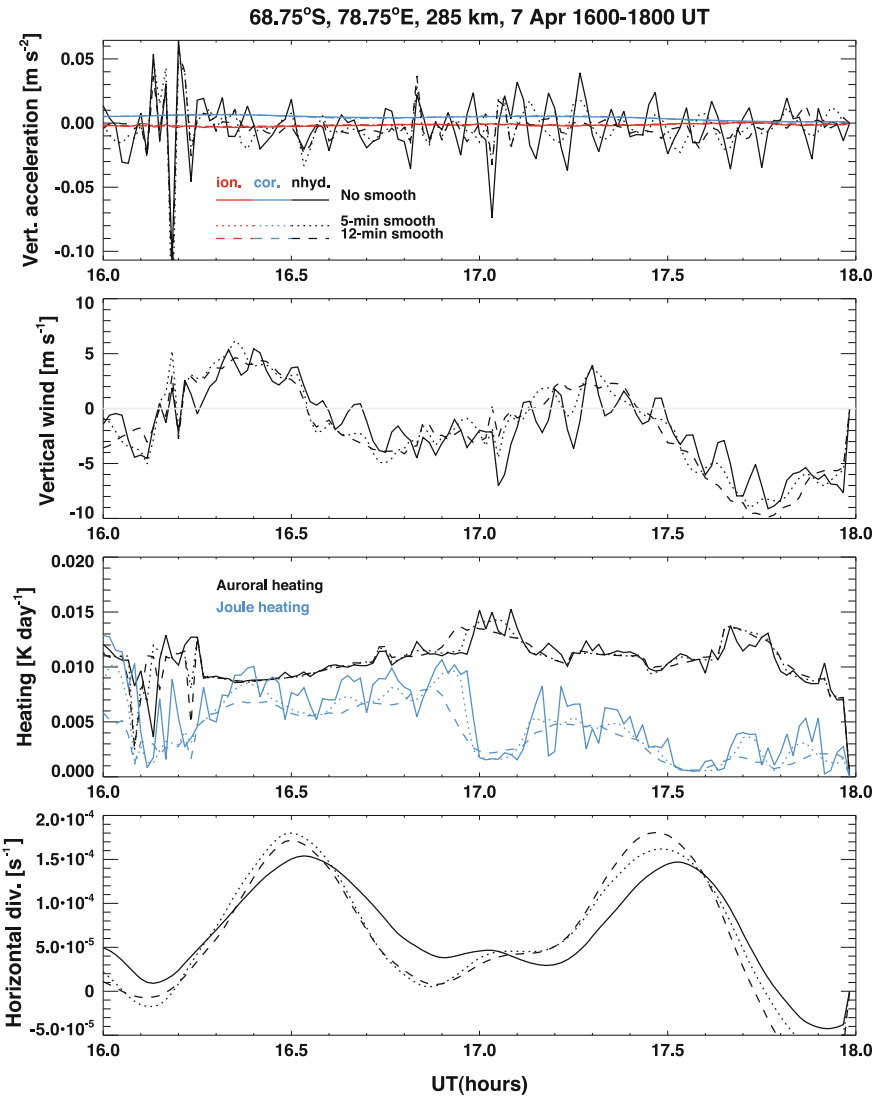

Fig. 5.8 Variations of vertical acceleration, vertical winds, Joule and auroral heating, and horizontal wind divergence at a representative Northern Hemisphere auroral latitude during April as simulated by the nonhydrostatic Gobal Ionosphere Thermosphere Model. (Yiğit et al. 2012b, Fig. 4)

wind can be understood as a plasma cloud that is being emitted radially outward in all directions. It carries high-energy charged particles that can reach Earth and cause disturbances in the geomagnetic field.

Typically magnetohydrodynamic (MHD) simulations are performed in order to predict the evolution of the solar wind, which originates at Sun and can propagate

Table 5.1 Conversion between the K_p and A_p indices

K_p	0o	0+	1–	1o	1+	2–	2o	2+	3–	3	3+	4–	4o	4+
A_p	0	2	3	4	5	6	7	9	12	15	18	22	27	32
K_p	5–	5o	5+	6–	6o	6+	7–	7o	7+	8–	8o	8+	9–	9o
A_p	39	48	56	67	80	94	111	132	154	179	207	236	300	400

toward Earth. One of the improved versions of the MHD models highlighted as a data-driven model have been presented in the work by Feng et al. (2015). The essence of the MHD models that are used for solar wind studies is to predict the changes of an electrically conducting solar wind, which is treated as a plasma fluid, in the presence of an ambient magnetic field. Hence, Maxwell's equations coupled with fluid dynamical conservation equations are solved. This recent model accurately predicted the solar wind behavior in comparison with the observed solar wind features during July/August 2008 period. However, It was less successful in predicting the total magnetic field strength during this period.

It is important to quantify geomagnetic activity in order to better predict space weather effects on the atmosphere. Various parameters/indices are used for this purpose, such as, Dst, K_p, A_p, IMF, etc., are studied to assess geomagnetic activity. The K_p index, which was introduced more than half a century ago (Bartels 1949), is by far the most commonly used index used by upper atmosphere scientists (Rostoker 1972). It has a scale of 0 to 9 determined at the end of 3-hour intervals as the maximum deviation of the magnetic field from the observed field as detected by a global network of magnetograms. Combination of all indices from different stations gives the planetary K_p index. Table 5.1 shows the conversion table between the A_p and K_p indices. Quiet geomagnetic conditions are characterized by K_p index of 0o–2+ ($A_p \sim$ 0–9), while only super storms reach K_p values larger than 8+.

5.6 Geomagnetic Storm Effects on the Upper Atmosphere

Let us next focus on one of the most prominent manifestations of space weather effects—geomagnetic storms. Magnetic storms are long-lasting and strong coupling of plasma from the solar wind into the magnetosphere. Following this coupling significant amount of energy and momentum is transferred to the thermosphere-ionosphere system. This energy and momentum transfer is a highly variable process and various magnetic parameters are observed in order to describe the temporal evolution of a storm. So, it is fair to say that magnetic storms are a particular manifestation of space weather in the geospace environment.

Temporal development of a storm is best seen in Dst. Figure 5.9 shows the storm-time Dst variation from 3 to 7 August 2011 as retrieved from the World Data Center for Geomagnetism, Kyoto. Around 1800 UT on 6 August a sudden disturbance of Dst is seen, marking the initial phase of the storm, which continuously grows in

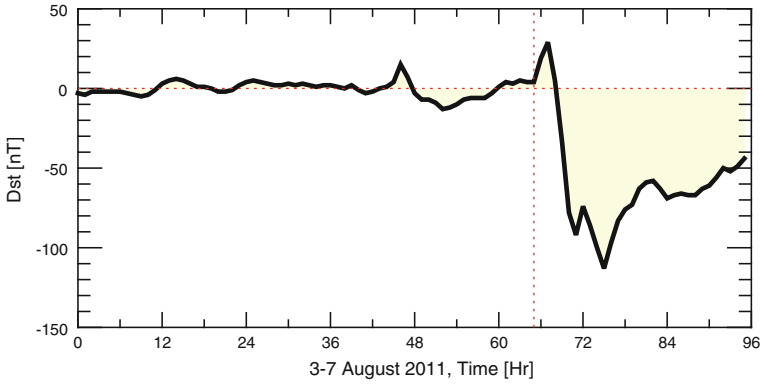

Fig. 5.9 Storm-time Dst variation from 3–7 August 2011. The horizontal and vertical *dotted red lines* denote the 0 nT level and the onset of the major magnetic disturbances, respectively. (Yiğit et al. 2016a, Fig. 1)

magnitude for about 6–8 h, indicating the main phase of the storm, after which Dst starts returning back to normal during the recovery phase that often lasts for several days. Moderate storms are characterized by $-50 \leq Dst \leq -100$ nT, while intense storms occur when $Dst < -100$ nT.

In the initial phase, increased solar wind speed v_x pushes the magnetopause inward ultimately compressing the magnetosphere. During the main phase, the induced changes in the magnetosphere (e.g., ring current enhancement) leads to a decrease of the geomagnetic field as measured on the ground. Overall, the magnetized solar wind interact with the geomagnetic field, producing substantial changes in particular at high-latitudes.

So, in upper atmospheric studies the magnetosphere is treated as a region of energy and momentum source. The flux of energy from the magnetosphere to the upper atmosphere can be quantified, for example, by the Poynting flux, which has been discussed in Sect. 2.7 and has been used in earlier studies as a diagnostic quantity (Thayer and Semester 2004; Knipp et al. 2011). A significant portion of the transferred energy manifests itself as a cross polar cap potential drop and to a minor extent as a particle precipitation around the auroral oval at high-latitudes (Burns et al. 2007). This deposited energy is then dissipated in the atmosphere leading to neutral gas heating by Joule heating and to variations in the radiative processes.

Owing to the nature of complex variations of the upper atmospheric response to geomagnetic storm, global-scale models have often been used to study how some upper atmospheric processes or variables respond to a storm. Using a simplified substorm model, Richmond and Matsushita (1975) revealed that large-scale gravity waves are launched at high-latitudes during the storm and they propagate equatorward with a speed of 750 m s^{-1}. These large phase speeds are characteristic of large-scale gravity waves. Earlier global-scale models such as CTIM (Fuller-Rowell et al. 1987) and TIGCM (Roble et al. 1988) have been used to study various aspects

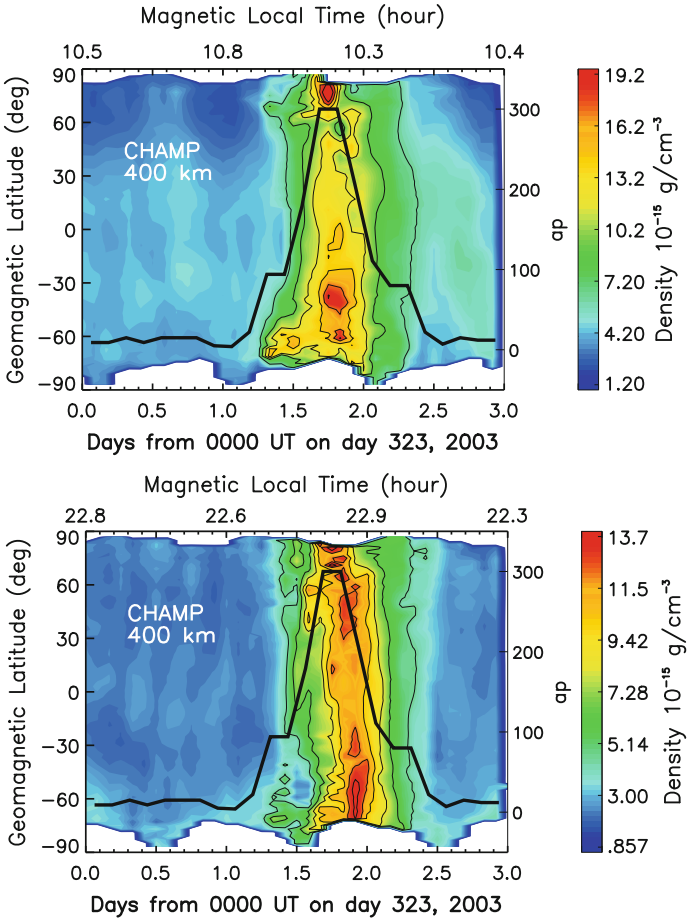

Fig. 5.10 Geomagnetic latitude-universal time (UT) distribution of the neutral density [g cm^{-3}] inferred from the CHAMP satellite during day-time (*top*) and night-time (*bottom*) as adopted from the work by Bruinsma et al. (2006), Fig. 3). The *black solid lines* represent the A_p index. Geomagnetic local times are specified above the figure

of storm-time thermospheric variations. Fuller-Rowell et al. (1994) have performed idealized GCM experiments in order to investigate the UT dependence in the thermospheric and ionospheric response to magnetospheric input. They showed that the thermosphere responds in a complex manner. Joule heating is enhanced, divergence horizontal winds drive vertical winds, which generate compositional changes at a fixed pressure level. Buonsanto et al. (1999) studied traveling atmospheric disturbances propagating equatorward from the high-latitudes by analyzing the neutral meridional wind perturbations modeled by a thermospheric GCM. Balan et al. (2012)'s combined analysis of the CHAMP, ROCSAT, and DMSP data showed that

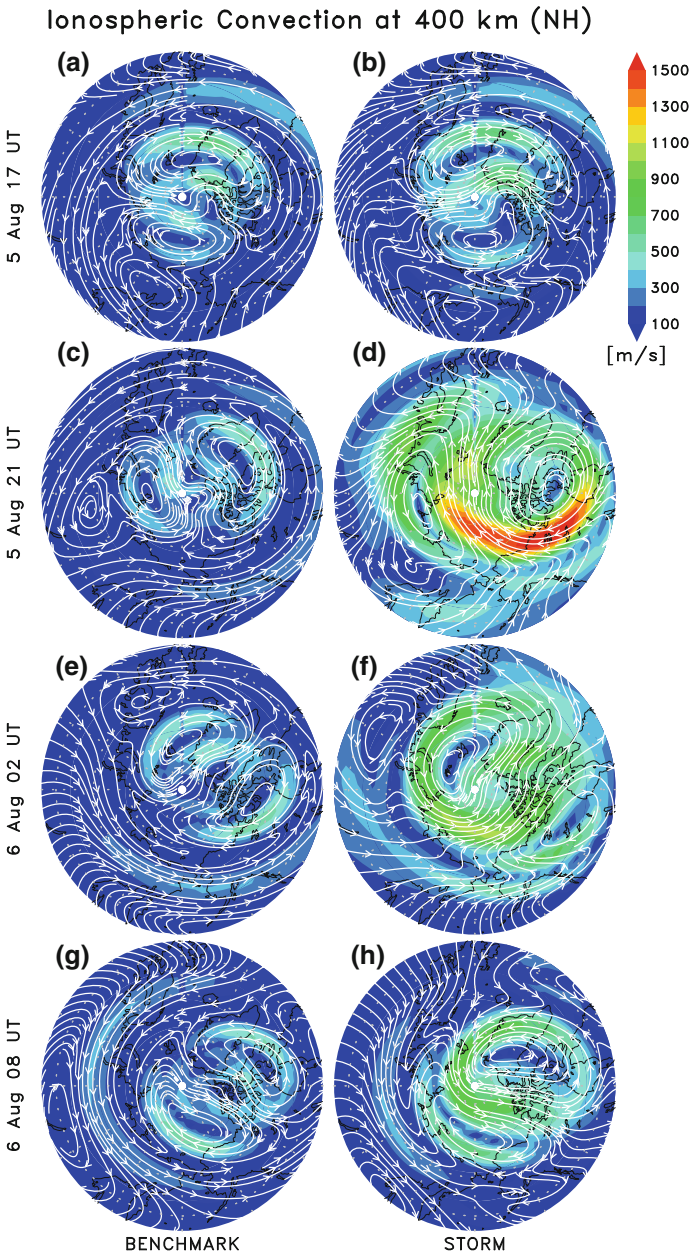

Fig. 5.11 Ionospheric plasma convection pattern at 400 km in the Northern Hemisphere simulated by the Global Ionosphere Thermosphere Model (GITM) during the different phases of the August 2011 major geomagnetic storm. Rows 1 to 4 show four representative UTs, 5 August 17 UT, 5 August 21 UT, 6 August 02 UT and 6 UT, respectively. The panels on the left shows the benchmark run, which is quiet-time run with constant low IMF, while panels on the right represent storm-time run with real-time IMF and auroral input. Adopted from Yiğit et al. (2016a, Fig. 5)

the equatorial thermosphere responds to a storm within 1.5 to 3 hours with respect to the onset of the main phase of the storm.

Bruinsma et al. (2006) have inferred from accelerometer measurements on board the CHAMP and GRACE satellites thermospheric density response to the sever geomagnetic storm of 20–21 November. During the main phase of the storm Dst has reached a value around –450 nT and and B_z dropped to –50 nT. Their results overall showed that density increased up to a factor of 3 to 8 in a global sense during the storm, which is shown in Fig. 5.10, where the top panel shows the day-time and bottom panel is for the night-time observations.

Recently, Yiğit et al. (2016a) have simulated with a GCM the response of thermosphere-ionosphere to the August 2011 geomagnetic storm, whose time evolution is illustrated in Fig. 5.9 by the Dst index, and studied the evolution of the ionospheric convection and hemispheric differences with respect to a quiet-time benchmark run. Figure 5.11 exemplifies their Northern Hemisphere results at 400 km, where the benchmark results are on the left column and the storm simulation is on the right. The top row shows the ion convection right before the storm onset and rows below show the evolution of the storm. It is seen that the convection pattern expands during the storm and the magnitude of ion convection intensifies significantly reaching 1500 m s^{-1}. Such large ion speeds have been seen in previous simulations of geomagnetic storms (Crowley et al. 1989) and are a significant source of ion drag and Joule heating for the neutrals.

5.7 Concluding Remarks

It has been nearly more than a hundred years since humanity has discovered the existence of a conducting layer at greater heights in the atmosphere, which we commonly refer to as the ionosphere. If you imagine the highest altitude a human reaches in the atmosphere is around ∼12 km while flying in an Airbus 380, of course unless we are speaking of an astronaut, a space tourist or a jet plane pilot. Although the upper atmosphere is far away from the surface, it has immense significance for satellite technology, navigation, telecommunication, planetary science missions, long-term climate change, and fundamental science.

With the advent of satellites, remote-sensing has become an integral part of observational research. In particular, advancements in satellite technology, communication and navigation systems, and in planetary science missions have highlighted the practical and technological relevance of the atmosphere-ionosphere system, which is a complex two-way nonlinearly coupled system of plasma and neutrals. Besides this complexity, we have increasingly learned that meteorological processes can appreciably contribute to the energy and momentum budget of the upper atmosphere. Various missions to other Solar System planets also suggest that planetary upper atmospheres are regions of interactions between plasma and neutrals under the influence of meteorological influences and space weather effects. Therefore, future studies of atmospheric and space physics have to increasingly adopt a whole atmosphere

approach coupled to lower atmospheric processes and effects originating at Sun. I envision that for future generation scientists it will be crucial to gain a solid knowledge in both atmospheric and space sciences in order to be able to deal with increasing degree of complexity in scientific and technological problems.

References

Baker JBH, Zhang Y, Greenwald RA, Paxton LJ, Morrison D (2004) Height-integrated joule and auroral particle heating in the night side high latitude thermosphere. Geophys Res Lett 31:L09807. doi:10.1029/2004GL019535

Balan N, Yamamoto M, nad Y Otsuka JYL, Liu H, Lühr H, (2011) New aspects of thermospheric and ionospheric storms revealed by CHAMP. J Geophys Res 116:A06320. doi:10.1029/2010JA016399

Balan N, Liu JY, Otsuka Y, Ram ST, Lühr H (2012) Ionospheric and thermospheric storms at equatorial latitudes observed by CHAMP, ROCSAT, and DMSP. J Atmos Sci 117:A01313. doi:10.1029/2011JA016903

Banks PM, Kockarts G (1973) Aeronomy. Part B. Academic Press, New York

Bartels J (1949) The standardized index, ks, and the planetary index, kp. IATME Bull 12b 97:2010–2021

Bates DR (1951) The temperature of the upper atmosphere. Proc Phys Soc 64B(9):805–821

Becker E (2004) Direct heating rates associated with gravity wave saturation. J Atmos Sol-Terr Phys 66:683–696

Becker E (2011) Dynamical control of the middle atmosphere. Space Sci Rev 168:283–314. doi:10.1007/s11214-011-9841-5

Becker E, Knpfel R, Lbken FJ (2015) Dynamically induced hemispheric differences in the seasonal cycle of the summer polar mesopause. J Atmos Sol-Terr Phys 129:128–141. doi:10.1016/j.jastp.2015.04.014, http://www.sciencedirect.com/science/article/pii/S1364682615000887

Beres JH, Alexander MJ, Holton JR (2004) A method of specifying the gravity wave spectrum above convection based on latent heating properties and background wind. J Atmos Sci 61(3):324–337. doi:10.1175/1520-0469(2004)061

Berger U, von Zahn U (1999) The two-level structure of the mesopause: a model study. J Geophys Res 104:22,083–22,093

Blanch E, Marsal S, Segarra A, Torta JM, Altadill D, Curto JJ (2013) Space weather effects on earths environment associated to the 24–25 October 2011 geomagnetic storm. Space Weather 11:153–168. doi:10.1002/swe.20035

Bruinsma S, Forbes JM, Nerem RS, Zhang X (2006) Thermosphere density response to the 20–21 November 2003 solar and geomagnetic storm from champ and grace accelerometer data. J Geophys Res Space Phys 111(A6)

Buonsanto M, Gonzalez S, Lu G, Reinisch B, Thayer J (1999) Coordinated incoherent scatter radar study of the January 1997 storm. J Geophys Res Space Phys 104(A11):24625–24637

Burns A, Solomon S, Wang W, Killeen T (2007) The ionospheric and thermospheric response to cmes: Challenges and successes. J Atmos Sol-Terr Phys 69(1):77–85

Canziani PO, Holton JR (1998) Kelvin waves and the quasi-biennial oscillation: An observational analysis. J Geophys Res 103:31509–31521

Chun HY, Baik JJ (2002) An updated parameterization of convectively forced gravity wave drag for use in large-scale models. J Geophys Res 59(5):1006–1017. doi:10.1175/1520-0469(2002)059

Clilverd MA, Rodger C, Andersson M, Seppälä A, Verronen P (2015) Linkages between the radiation belts, polar atmosphere and climate: electron precipitation through wave particle interactions. in waves, particles and storms in geospace. In: Balasis G, Daglis IA, Mann IR (eds) Waves, particles and storms in geospace. Oxford University Press

Crowley G, Emery BA, Roble RG, Carlson HC, Knipp DJ (1989) Thermospheric dynamics during September 18–19, 1984, 1. model simulations. J Geophys Res 94(A12):16925–16944

Daniell RE, Brown LD, Anderson DN, Fox MW, Doherty PH, Decker DT, Sojka JJ, Schunk RW (1995) Pim: a global ionospheric parameterization based on first principles models. R Sci 30:1499–1510

Deng Y, Maute A, Richmond AD, Roble RG (2009) Impact of electric field variability on joule heating and thermospheric temperature and density. Geophys Res Lett 36:L08105. doi:10.1029/2008GL036916

Dickinson RE, Roble RG, Ridley EC (1971) Response of the neutral thermosphere at f-layer heights to interaction of a global wind with anomalies of ionozation. J Atmos Sci 28:1280–1293

Dickinson RE, Ridley EC, Roble RG (1981) A three-dimensional general circulation model of the thermosphere. J Geophys Res 86(A3):1499–1512

Fagundes PR, Goncharenko LP, de Abreu AJ, Venkatesh K, Pezzopane M, de Jesus R, Gende M, Coster AJ, Pillat VG (2015) Ionospheric response to the 2009 sudden stratospheric warming over the equatorial, low, and middle latitudes in the South American sector. J Geophys Res Space Phys 120(9):7889–7902. doi:10.1002/2014JA020649

Feng X, Ma X, Xiang C (2015) Data-driven modeling of the solar wind from 1 rs to 1 au. J Geophys Res Space Phys 120(12):10159–10174. doi:10.1002/2015JA021911

Fomichev VI, Blanchet JP, Turner DS (1998) Matrix parameterization of the 15 μm band cooling in the middle atmosphere for variable CO_2 concentration. J Geophys Res 103:D10, 11505–11528

Fuller-Rowell T, Rees D, Quegan S, Moffett R, Bailey G (1987) Interactions between neutral thermospheric composition and the polar ionosphere using a coupled ionosphere-thermosphere model. J Geophys Res Space Phys 92(A7):7744–7748

Fuller-Rowell T, Codrescu M, Moffett R, Quegan S (1994) Response of the thermosphere and ionosphere to geomagnetic storms. J Geophys Res Space Phys 99(A3):3893–3914

Fuller-Rowell TJ (1981) A three-dimensional, time-dependent model of the thermosphere. PhD thesis, University of London

Fuller-Rowell TJ, Rees D (1981) A three-dimensional, time-dependent simulation of the global dynamical response of the thermosphere to a geomagnetic storm. J Atmos Terr Phys 43(7):701–721

Galand M, Fuller-Rowell TJ, Codrescu MV (2001) Response of the upper atmosphere auroral protons. J Geophys Res 106(A1):127–139

Galvan DA, Komjathy A, Hickey MP, Stephens P, Snively J, Tony Song Y, Butala MD, Mannucci AJ (2012) Ionospheric signatures of tohoku-oki tsunami of March 11, 2011: Model comparisons near the epicenter. R Sci 47(4). doi:10.1029/2012RS005023

Garcia RR, Solomon S (1985) The effect of breaking gravity waves on the dynamics and chemical composition of the mesosphere and lower thermosphere. J Geophys Res 90:3850–3868, implementation of Lindzen's parameterization into a two-dimensional dynamical model to study the effects of GWs in the MLT

Garcia RR, Marsh DR, Kinnison DE, Boville BA, Sassi F (2007) Simulations of secular trends in the middle atmosphere. J Geophys Res 112:D09301. doi:10.1029/2006JD007485

Goncharenko LP, Hsu VW, Brum CGM, Zhang SR, Fentzke JT (2013) Wave signatures in the midlatitude ionosphere during a sudden stratospheric warming of January 2010. J Geophys Res Space Phys 118. doi:10.1029/2012JA018251

Gong Y, Zhou Q, Zhang SD, Aponte N, Sulzer M, Gonzalez S (2013) The F region and topside ionosphere response to a strong geomagnetic storm at Arecibo. J Geophys Res 118:51775183. doi:10.1002/jgra.50502

Guharay A, Batista P, Clemesha B (2015) On the variability of the diurnal tide and coupling with planetary waves in the MLT over cachoeira paulista (22.7s, 45w). J Atmos Sol-Terr Phys 133:7–17. doi:10.1016/j.jastp.2015.07.016

Hickey MP, Huang TY, Walterscheid R (2003) Gravity wave packet effects on chemical exothermic heating in the mesopause region. J Geophys Res: Space Phys 108(A12):1448. doi:10.1029/2002JA009363

Hickey MP, Schubert G, Walterscheid RL (2009) Propagation of tsunami-driven gravity waves into the thermosphere and ionosphere. J Geophys Res. doi:10.1029/2009JA014105
Hickey MP, Walterscheid RL, Schubert G (2010) Wave mean flow interactions in the thermosphere induced by a major tsunami. J Geophys Res Space Phys 115(A9). doi:10.1029/2009JA014927
Holton JR (1982) The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere. J Atmos Sci 39:791–799
Holton JR, Lindzen RS (1972) An updated theory for the quasi-biennial cycle of the tropical stratosphere. J Atmos Sci 29:1076–1080
Huang CY, Su YJ, Sutton EK, Weimer DR, Davidson RL (2014) Energy coupling during the august 2011 magnetic storm. J Geophys Res Space Phys 119. doi:10.1002/2013JA019297
Kazimirovsky E, Herraiz M, Morena BADL (2003) Effects on the ionosphere due to phenomena occurring below it. Surv Geophys 24:139–184
Kellogg WW (1961) Chemical heating above the polar mesopause in winter. J Meteorol 18(3):373–381. doi:10.1175/1520-0469 (1961) 018⟨0373:CHATPM⟩2.0.CO;2
Klimenko M, Klimenko V (2012) Disturbance dynamo, prompt penetration electric field and over-shielding in the earth's ionosphere during geomagnetic storm. J Atmos Sol-Terr Phys 9091:146–155. doi:10.1016/j.jastp.2012.02.018, recent Progress in the Vertical Coupling in the Atmosphere-Ionosphere System
Knipp D, Eriksson S, Kilcommons L, Crowley G, Lei J, Hairston M, Drake K (2011) Extreme poynting flux in the dayside thermosphere: Examples and statistics. Geophys Res Lett 38(16). doi:10.1029/2011GL048302
Knipp DJ, Tobiska WK, Emery BA (2004) Direct and indirect thermospheric heating sources for solar cycles 21–23. Solar Phys 224(1):495. doi:10.1007/s11207-005-6393-4
Knizova PK, Kouba D, K Potužnikova JB (2015) Influence of meteorological systems on the ionosphere over europe. J Atmos Sol-Terr Phys
Kockarts G (1980) Nitric oxide cooling in the terrestrial thermosphere. Geophys Res Lett 7:137–140
Kuroda T, Medvedev AS, Yiğit E, Hartogh P (2015) A global view of gravity waves in the martian atmosphere inferred from a high-resolution general circulation model. Geophys Res Lett. doi:10.1002/2015GL066332
Kwak YS, Richmond AD (2007) An analysis of the momentum forcing in the high-latitude lower thermosphere. J Geophys Res 112(A0):1306. doi:10.1029/2006JA011910
Laskar FI, Pallamraju D, Veenadhari B, Vijaya Lakshmi T, Anji Reddy M, Chakrabarti S (2015) Gravity waves in the thermosphere: solar activity dependence. Adv Space Res 55:1651–1659. doi:10.1016/j.asr.2014.12.040
Laštovička J (2006) Forcing of the ionosphere by waves from below. J Atmos Sol-Terr Phys 68:479–497
Lübken FJ (1999) Thermal structure of the arctic summer mesosphere. J Geophys Res: Atmos 104(D8):9135–9149. doi:10.1029/1999JD900076
Matsuo T, Richmond AD (2008) Effects of high-latitude ionospheric electric field variability on global thermospheric Joule heating and mechanical energy transfer rate. J Geophys Res 113:A07309. doi:10.1029/2007JA012993
Medvedev AS, Klaassen GP (2000) Parameterization of gravity wave momentum deposition based on nonlinear wave interactions: Basic formulation and sensitivity tests. J Atmos Sol-Terr Phys 62:1015–1033
Medvedev AS, Klaassen GP (2003) Thermal effects of saturating gravity waves in the atmosphere. J Geophys Res 108(D2):4040. doi:10.1029/2002JD002504
Medvedev AS, Yiğit E, Hartogh P, Becker E (2011) Influence of gravity waves on the Martian atmosphere: general circulation modeling. J Geophys Res 116:E10004. doi:10.1029/2011JE003848
Medvedev AS, Yiğit E, Kuroda T, Hartogh P (2013) General circulation modeling of the martian upper atmosphere during global dust storms. J Geophys Res Planets 118:1–13. doi:10.1002/jgre.20163

Medvedev AS, González-Galindo F, Yiğit E, Feofilov AG, Forget F, Hartogh P (2015) Cooling of the martian thermosphere by CO_2 radiation and gravity waves: an intercomparison study with two general circulation models. J Geophys Res Planets 120. doi:10.1002/2015JE004802

Medvedev AS, Nakagawa H, Mockel C, Yiğit E, Kuroda T, Hartogh P, Terada K, Terada N, Seki K, Schneider NM, Jain SK, Evans JS, Deighan JI, McClintock WE, Lo D, Jakosky BM (2016) Comparison of the martian thermospheric density and temperature from iuvs/maven data and general circulation modeling. Geophys Res Lett 43(7):3095–3104. doi:10.1002/2016GL068388

Mikhailov AV (2008) Ionospheric f1 layer long-term trends and the geomagnetic control concept. Ann Geophys 26:3793–3803

Miyoshi Y, Fujiwara H, Jin H, Shinagawa H (2014) A global view of gravity waves in the thermosphere simulated by a general circulation model. J Geophys Res Space Phys 119:5807–5820. doi:10.1002/2014JA019848

Miyoshi Y, Fujiwara H, Jin H, Shinagawa H (2015) Impacts of sudden stratospheric warming on general circulation of the thermosphere. J Geophys Res: Space Phys 120(12):10897–10912. doi:10.1002/2015JA021894

Namgaladze A, Korenkov YN, Klimenko V, Karpov I, Bessarab F, Surotkin V, Glushchenko T, Naumova N (1988) Global model of the thermosphere-ionosphere-protonosphere system. In: Ionospheric modelling, Springer, pp 219–254

Nisbet JS (1975) Models of the ionosphere. In: McCormac BM (ed) Atmospheres and Earth and the planets, D. Reidel publishing company, vol 51, pp 243–258

Oberheide J, Shiokawa K, Gurubaran S, Ward WE, Fujiwara H, Kosch MJ, Makela JJ, Takahashi H (2015) The geospace response to variable inputs from the lower atmosphere: a review of the progress made by Task Group 4 of CAWSES-II. Progr Earth Planet Sci 2. doi:10.1186/s40645-014-0031-4

Pancheva D, Mukhtarov P, Andonov B, Forbes J (2010) Global distribution and climatological features of the 56-day planetary waves seen in the saber/timed temperatures (20022007). J Atmos Sol-Terr Phys 72(1):26–37. doi:10.1016/j.jastp.2009.10.005, http://www.sciencedirect.com/science/article/pii/S1364682609002612

Pancheva D, Miyoshi Y, Mukhtarov P, Jin H, Shinagawa H, Fujiwara H (2012) Global response of the ionosphere to atmospheric tides forced from below: Comparison between cosmic measurements and simulations by atmosphere-ionosphere coupled model gaia. J Geophys Res Space Phys 117(A7). doi:10.1029/2011JA017452

Pulkkinen A, Bernabeu E, Thomson A, Viljanen A, Pirjola R, Boteler D, Eichner J, Cilliers PJ, Welling D, Savani NP, Weigel RS, Love JJ, Balch C, Ngwira CM, Crowley G, Schultz A, Kataoka R, Anderson B, Fugate D, Simpson JJ, MacAlester M (2017) Geomagnetically induced currents: Science, engineering and applications readiness. Space Weather pp n/a–n/a. doi:10.1002/2016SW001501

Richardson LF (1922) Weather prediction by numerical process. Cambridge

Richmond AD, Matsushita S (1975) Thermospheric response to a magnetic substorm. J Geophys Res 80(19):2839–2850

Rienecker MM, Coauthors (2008) The GEOS-5 data assimilation system – documentation of versions 5.0.1 and 5.1.0, and 5.2.0. NASA tech. rep. series on global modeling and data assimilation. Tech. Rep. NASA/TM-2008-104606, NASA

Rishbeth H (2002) High above the earth. Nature 418

Roble RG, Emery BA (1983) On the global mean temperature of the thermosphere. Planet Space Sci 31:597–614

Roble RG, Dickinson RE, Ridley EC (1977) Seasonal and solar cycle variations of the zonal mean circulation in the thermosphere. J Geophys Res 82:5493–5504

Roble RG, Ridley EC, Richmond AD (1988) A coupled thermosphere/ionosphere general circulation model. Geophys Res Lett 15(12):1325–1328

Rostoker G (1972) Geomagnetic indices. Rev Geophys 10(4):935–950

Simmons AJ, Burridge DM (1981) An energy and angular momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. Mon Weather Rev 109:758–766

Sreeja V (2016) Impact and mitigation of space weather effects on gnss receiver performance. Geosci Lett 3(24). doi:10.1186/s40562-016-0057-0

Thayer JP, Semester J (2004) The convergence of magnetospheric energy flux in the polar atmosphere. J Atmos Sol-Terr Phys 66:807–824

Tonev PT, Velinov PIY (2015) Vertical coupling between troposphere and lower ionosphere by electric currents and fields at equatorial latitudes. J Atmos Sol-Terr Phys. doi:10.1016/j.jastp.2015.10.012

Vadas SL, Taylor MJ, Pautet PD, Stamus PA, Fritts DC, Liu HL, Sabbas FTS, Rampinelli VT, Batista P, Takahashi H (2009) Convection: the likely source of the medium-scale gravity waves observed in the oh airglow layer near brasilia, brazil, during the spreadfex campaign. Ann Geophys 27:231–259

VanZandt TE (1982) A universal spectrum of buoyancy waves in the atmosphere. Geophys Res Lett 9:575–578

Walterscheid RL, Hickey MP, Schubert G (2013) Wave heating and jeans escape in the martian upper atmosphere. J Geophys Res Planets 118(2413–2422):2169–9402. doi:10.1002/jgre.20164

Weinstock J (1982) Nonlinear theory of gravity waves: Momentum deposition, generalized rayleigh friction, and diffusion. J Atmos Sci 39:1698–1710

Wilson GR, Weimer DR, Wise JO, Marcos FA (2006) Response of the thermosphere to Joule heating and particle precipitation. J Geophys Res A10314. doi:10.1029/2005JA011274

Yiğit E (2015) Atmospheric and space sciences: neutral atmospheres, vol 1. SpringerBriefs in Earth Science. Springer, Netherlands. doi:10.1007/978-3-319-21581-5

Yiğit E, Medvedev AS (2009) Heating and cooling of the thermosphere by internal gravity waves. Geophys Res Lett 36:L14807. doi:10.1029/2009GL038507

Yiğit E, Medvedev AS (2012) Gravity waves in the thermosphere during a sudden stratospheric warming. Geophys Res Lett 39:L21101. doi:10.1029/2012GL053812

Yiğit E, Medvedev AS (2013) Extending the parameterization of gravity waves into the thermosphere and modeling their effects. In: Lübken FJ (ed) Climate and weather of the Sun-Earth system (CAWSES). Springer Atmospheric Sciences. Springer, Netherlands, pp 467–480. doi:10.1007/978-94-007-4348-9_25

Yiğit E, Medvedev AS (2015) Internal wave coupling processes in Earth's atmosphere. Adv Space Res 55(5):983–1003. doi:10.1016/j.asr.2014.11.020, http://www.sciencedirect.com/science/article/pii/S0273117714007236

Yiğit E, Medvedev AS (2016) Role of gravity waves in vertical coupling during sudden stratospheric warmings. Geosci Lett 3:27. doi:10.1186/s40562-016-0056-1

Yiğit E, Aylward AD, Medvedev AS (2008) Parameterization of the effects of vertically propagating gravity waves for thermosphere general circulation models: Sensitivity study. J Geophys Res 113:D19106. doi:10.1029/2008JD010135

Yiğit E, Medvedev AS, Aylward AD, Hartogh P, Harris MJ (2009) Modeling the effects of gravity wave momentum deposition on the general circulation above the turbopause. J Geophys Res 114:D07101. doi:10.1029/2008JD011132

Yiğit E, Medvedev AS, Aylward AD, Ridley AJ, Harris MJ, Moldwin MB, Hartogh P (2012a) Dynamical effects of internal gravity waves in the equinoctial thermosphere. J Atmos Sol-Terr Phys 90–91:104–116. doi:10.1016/j.jastp.2011.11.014

Yiğit E, Ridley AJ, Moldwin MB (2012b) Importance of capturing heliospheric variability for studies of thermospheric vertical winds. J Geophys Res 117:A07306. doi:10.1029/2012JA017596

Yiğit E, Medvedev AS, England SL, Immel TJ (2014) Simulated variability of the high-latitude thermosphere induced by small-scale gravity waves during a sudden stratospheric warming. J Geophys Res Space Phys 119. doi:10.1002/2013JA019283

Yiğit E, Frey HU, Moldwin MB, Immel TJ, Ridley AJ (2016a) Hemispheric differences in the response of the upper atmosphere to the August 2011 geomagnetic storm: a simulation study. J Atmos Sol-Terr Phys 141:13–26. doi:10.1016/j.jastp.2015.10.002

Yiğit E, Knížová PK, Georgieva K, Ward W (2016b) A review of vertical coupling in the atmosphere-ionosphere system: Effects of waves, sudden stratospheric warmings, space weather, and of

solar activity. J Atmos Sol-Terr Phys 141:1–12. doi:10.1016/j.jastp.2016.02.011, http://www.sciencedirect.com/science/article/pii/S1364682616300426
Zülicke C, Becker E (2013) The structure of the mesosphere during sudden stratospheric warmings in a global circulation model. J Geophys Res Atmos 118. doi:10.1002/jgrd.50219

Erratum to: Atmospheric and Space Sciences: Ionospheres and Plasma Environments

Erratum to:
E. Yigit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*, SpringerBriefs in Earth Sciences, https://doi.org/10.1007/978-3-319-62006-0

The original version of the book was inadvertently published with incorrect author name "E. Yigit" which has to be corrected as "E. Yiğit" in source line of all chapters. The erratum book has been updated with the change

The updated online version of this book can be found at https://doi.org/10.1007/978-3-319-62006-0

E. Yiğit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*, SpringerBriefs in Earth Sciences; https://doi.org/10.1007/978-3-319-62006-0_6

Appendix A
Physical Constants and Parameters

Physical Constants

In Table A.2 the physical constants that are used in this book are summarized (Table A.1).

Table A.1 List of physical constants and units used in the book in alphabetical order. Approximate values of the constants are stated

Constant	Label	Value	Dimension
Astronomical unit	AU	149.6×10^{6}	km
Atomic mass unit	u	1.66×10^{-27}	kg
Avogadro's number	N_A	6.02×10^{23}	mol^{-1}
Boltzmann constant	k_b	1.3807×10^{-23}	J K^{-1}
Electron mass	m_e	9.109×10^{-31}	kg
Electron volt	eV	1.6022×10^{-19}	J
Elementary charge	e	1.602×10^{-19}	C
Proton mass	m_p	1.672×10^{-27}	kg
Proton/electron mass ratio	m_p/m_e	1836.15	
Gravitational constant	G	6.67×10^{-11}	$\text{N m}^2\ \text{kg}^{-2}$
Permittivity of free space	$\varepsilon_0 = 1/(\mu_0 c^2)$	8.854×10^{-12}	F m^{-1}
Permeability of free space	$\mu_0 = 1/(\varepsilon_0 c^2)$	$4\pi \times 10^{-7}$	H m^{-1}

(continued)

E. Yiğit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-62006-0

Table A.1 (continued)

Constant	Label	Value	Dimension
Planck constant	h	6.626×10^{-34}	J s
Solar constant	S_c	1368	W m^{-2}
Speed of light in vacuum	c	2.9979×10^{8}	m s^{-1}
Stefan-Boltzmann constant	σ_b	5.67×10^{-8}	W m^{-2} K^{-4}
Universal gas constant	R	8.317	J $mol^{-1}K^{-1}$

Table A.2 A brief list of useful unit conversions are listed below alphabetically

Field	Unit name	Unit abbreviation	Unit conversion
Capacitance	farad	F	C/V
Energy	joule	J	kg m^2 s^{-2}
Inductance	henry	H	kg m^2 s^{-2} A^{-2}

Some Useful Unit Conversions

Sometimes it is convenient to relate different units to each other.

Notation of Physical Parameters

Table A.3 summarized a list of major physical parameters used in the book.

Table A.3 A list of important physical parameters used in the book. This is not a full list of parameters

Notation	Parameter	Dimension
a	Acceleration	m s^{-2}
B	Magnetic field vector	T
$\mathbf{c}_s$	Species three-dimensional thermal velocity vector	m s^{-1}
c_p	Specific heat capacity at constant pressure	J K^{-1} kg^{-1}
c_v	Specific heat capacity at constant volume	J K^{-1} kg^{-1}
E	Electric field vector	V m^{-1} or N C^{-1}
E	Kinetic energy	J

(continued)

Table A.3 (continued)

Notation	Parameter	Dimension
f_s	Species distribution function	$m^{-3}\ (m\ s^{-1})^{-1}$
$f(\mathbf{r}, \mathbf{v}, t)$	Phase space distribution function	$m^{-3}\ (m\ s^{-1})^{-1}$
$\mathbf{F}_L$	Lorenz force	N
$\mathscr{F}_M$	Magnetic force density	$N\ m^{-3}$
$\mathscr{F}$	Force per unit mass	$N\ kg^{-1}$
H_p	Plasma scale height	m
H_s	Species scale height	m
$\mathbf{j}$	Current density (charge flux)	$A\ m^{-2}$
m	Molecular mass	kg
M	Molar mass	$kg\ mol^{-1}$
m_e	Electron mass	kg
m_i	Ion mass	kg
m_n	Neutral mass	kg
m_p	Proton mass	kg
n	Number density	m^{-3}
n_e	Electron number density	m^{-3}
n_i	Ion number density	m^{-3}
p	Pressure	$kg\ m^{-1}\ s^{-2} = Pa$
q_{uv}	Heat production due to solar UV radiation	$J\ m^3\ s^{-1}$
t	Time	s
T	Temperature	K
T_e	Electron temperature	K
T_i	Ion temperature	K
T_p	Plasma temperature	K
$\mathbf{u}$	Three-dimensional neutral drift velocity	$m\ s^{-1}$
$\mathbf{u}_i$	Three-dimensional ion drift velocity	$m\ s^{-1}$
u	Zonal (drift) speed	$m\ s^{-1}$
$\mathbf{v}$	Three-dimensional velocity vector	$m\ s^{-1}$
v	Meridional (drift) speed	$m\ s^{-1}$
$\mathbf{v}_s$	Three-dimensional species velocity vector	$m\ s^{-1}$
V	Volume	m^3
w	Vertical (drift) speed	$m\ s^{-1}$
$\mathscr{E}$	Energy deposition rate	$J\ m^3\ s^{-1}$
$\mathscr{E}_J$	Energy deposition rate due to Joule heating	$J\ m^3\ s^{-1}$
σ	Electrical conductivity	$Si\ m^{-1}$
σ_s	Absorption cross section of a species	m^2
σ	Stress tensor	$kg\ m^{-1}\ s^{-2}$
	Conductivity tensor	
κ	Thermal conductivity	$J\ K^{-1}\ m^{-1}\ s^{-1}$
κ_t	Turbulent thermal conductivity	$J\ K^{-1}\ m^{-1}\ s^{-1}$

(continued)

Table A.3 (continued)

Notation	Parameter	Dimension
κ_m	Molecular thermal conductivity	J K^{-1} m^{-1} s^{-1}
Σ_p	Height-integrated Pederson conductivity	S
Σ_H	Height-integrated Hall conductivity	S
ρ	Mass density	kg m^{-3}
ρ_c	Charge density	C m^{-3}
ρ_r	Electrical resistivity	Ω m
λ_D	Debye length	m
v	Specific volume ($= \rho^{-1}$)	m^3 kg^{-1}
ν_{ij}	Collision frequency of the i-th species with the j-th species	s^{-1}
ν_{in}	Ion-neutral collision frequency	s^{-1}
ν_{en}	Electron-neutral collision frequency	s^{-1}
ω	Angular frequency	rad s^{-1}
τ	Optical depth (thickness)	
τ	Shear stress tensor	kg m^{-1} s^{-2}

Appendix B
Mathematical Tools

B.1 Levi-Civita Symbol

The value of the Levi-Civita symbol ε_{ijk}, which is a third-order Cartesion tensor, is given by

$$\varepsilon_{ijk} = \begin{cases} 1 \text{ if } i,\ j, k \text{ is an even permutation of } 1, 2, 3 \\ -1 \text{ if } i,\ j, k \text{ is an odd permutation of } 1, 2, 3 \\ 0 \text{ otherwise} \end{cases} \tag{B.1}$$

For example, $\varepsilon_{ijk} = \varepsilon_{jki} = \varepsilon_{kij} = 1$.

B.2 Divergence Theorem

The divergence theorem relates a volume integral to an area integral for a continuous and differentiable vector field

$$\int_V \nabla \cdot \mathbf{a}\, \mathrm{d}V = \int_S \mathbf{a} \cdot \mathrm{d}\mathbf{S}. \tag{B.2}$$

For example, the divergence of an electric field $\mathbf{E} = \mathbf{F}/q$ can be expressed

$$\int_V \nabla \cdot \mathbf{E}\, \mathrm{d}V = \oint_S \mathbf{E} \cdot \mathrm{d}\mathbf{S} \tag{B.3}$$

E. Yiğit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-62006-0

The left-hand side represents the total source in the volume, while the right-hand side represents total flux through the surface, as defined as the surface integral over the boundary of the volume. In this specific example, the surface integral of the electric field is the net charge enclosed in the surface,

$$\oint_S \mathbf{E} \cdot \mathrm{d}\mathbf{S} = \frac{q}{\varepsilon_0}, \tag{B.4}$$

which is essentially a form of Maxwell's equation 2.30.

Glossary

Chapman Layers A simplified description of an ionospheric layer whose ion concentration demonstrates a solar zenith angle dependence. Transport processes, such as diffusion or wave-induced effects are neglected. Ionospheric E and F_1 layers are approximately Chapman-like but the F_2 layer considerably departs from a simple Chapman layer.

Charged Particles Particles that carry either a negative or a positive charge. Atomic particles like protons and electrons are charged particles. More complex atoms or molecules can become an ion, either a negative or a positive ion, when they undergo chemical or radiative reactions.

Diffusive Equilibrium In the upper atmosphere above about 100 km, the molecular diffusion dominates over turbulent diffusion, thus chemical species are diffusively separated depending on their molecular mass, hence on their species scale height. Lighter species have larger scale height, and thus they become more prevalent at higher altitudes in planetary thermospheres.

Electric Field A field that originate at a charged particle. Charged particles experience an electric force when they are placed in an electric field associated with another charge.

Fluid Theory The theoretical model of fluids, in which the individual particles perform a collective motion owing to mutual interactions and are thus described by a single temperature. Therefore, the individual behavior of the particles are not modeled.

General Circulation Models Time-dependent nonlinear numerical models of the atmosphere and/or ionosphere solving the fundamental conservation equations on a three-dimensional grid used to better understand the state and the evolution of the atmospheric environment.

Geomagnetic Storms Substantial disturbance of the geomagnetic field due to processes originating at Sun. The solar wind impinges on the magnetosphere, resulting in ring current enhancement, which is coupled to the atmosphere-ionosphere

E. Yiğit, *Atmospheric and Space Sciences: Ionospheres and Plasma Environments*,
SpringerBriefs in Earth Sciences, DOI 10.1007/978-3-319-62006-0

system. Thus this ring current enhancement manifest itself as a depression of the geomagnetic field as measured on the ground.

Ionogram An ionosonde recording that display the relationship between the frequency of the emitted signal and the virtual height of the reflected signal in order to derive various observational parameters of the ionosphere.

Ionosphere The partially ionized portion of a planetary atmosphere formed primarily by photoionization processes.

Ion drag In the upper atmosphere, neutral air motion is influenced by collisions with charged particle motion. This frictional effect of ions on the neutrals is called the ion drag and is important in the F region ionosphere of Earth.

Kinetic Theory In contrary to the fluid model, the kinetic model treats the individual species separately, assigning them a different velocity and temperature, denoted by $\mathbf{v}_s$ and T_s.

Magnetic Field A distribution of field in which electrically charged particles moving with a net perpendicular component to the (magnetic) field experience a force perpendicular both to the direction of their motion and the seed field.

Magnetohydrodynamics Application of the equations of fluid dynamics to the description of plasma as a fluid. Field equations are transport equations are typically both solved in a self-consistent manner.

Maxwell-Boltzmann Velocity Distribution is a velocity distribution function that describes systems in the equilibrium state. Various assumptions are made in order to simplify the molecular motion: (1) the molecular size is much smaller than typical distance between molecules; (2) collisions are rare; particles perform random motion; (3) intermolecular forces are neglected. Probability of encountering particles decreases exponentially as the energy increases.

Phase space Six-dimentional space in which a particle is described by three velocity coordinates (velocity-space) and three spatial coordinates (configuration space).

Photoionization Formation of an ion-electron pair by the absorption of solar radiation, typically at high frequencies.

Plasma A plasma is an electrically quasi-neutral gas that is mainly composed of (positively and negatively) charged particles that exhibit collective motion.

Solar constant The average amount of radiative solar energy (i.e., total energy flux) reaching the top of Earth's atmosphere at a mean distance of 1 AU ($=149.6\times10^6$ km).

Space weather A collective expression for the effects of Sun and its extended magnetic field environment on Earth's atmosphere and geospace environment.

Transport Equations Fundamental equations that describe the transport of momentum, energy, and mass in fluids.

Velocity Moments The procedure of multiplying the species distribution function f_s by powers or products of $\mathbf{v}_s$ and then integrating over all species velocities $\mathbf{v}_s$.

Vertical Coupling The interactions between the different atmospheric layers in the vertical direction generated for example by upward propagating acoustic waves, gravity waves, tides, and planetary waves.