

Ionospheres

Physics, Plasma Physics, and Chemistry

Robert W. Schunk
Utah State University

Andrew F. Nagy
University of Michigan



PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK <http://www.cup.cam.ac.uk>

40 West 20th Street, New York, NY 10011-4211, USA <http://www.cup.org>

10 Stamford Road, Oakleigh, Melbourne 3166, Australia

Ruiz de Alarcón 13, 28014 Madrid, Spain

© Robert W. Schunk and Andrew F. Nagy 2000

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2000

Printed in the United States of America

Typeface Times Roman 10 $\frac{1}{4}$ /13 $\frac{1}{2}$ pt. and Joanna *System* L^AT_EX 2_ε [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data

Schunk, R. W.

Ionospheres : physics, plasma physics, and chemistry / Robert W. Schunk, Andrew F. Nagy.

p. cm. – (Cambridge atmospheric and space science series)

ISBN 0-521-63237-4 (hb)

1. Ionosphere. 2. Planets – Ionospheres. I. Nagy, Andrew. II. Title. III. Series.

QC881.2.I6 S435 2000

538'.767'09992 – dc21

99-054707

ISBN 0 521 63237 4 hardback

Contents

Chapter 1 **Introduction** 1

- 1.1 Background and Purpose 1
- 1.2 History of Ionospheric Research 3
- 1.3 Specific References 8
- 1.4 General References 9

Chapter 2 **Space Environment** 11

- 2.1 Sun 11
- 2.2 Interplanetary Medium 16
- 2.3 Earth 21
- 2.4 Inner Planets 31
- 2.5 Outer Planets 36
- 2.6 Moons and Comets 38
- 2.7 Plasma and Neutral Parameters 41
- 2.8 Specific References 43

Chapter 3 **Transport Equations** 47

- 3.1 Boltzmann Equation 47
- 3.2 Moments of the Distribution Function 50
- 3.3 General Transport Equations 52
- 3.4 Maxwellian Velocity Distribution 55
- 3.5 Closing the System of Transport Equations 56

3.6	13-Moment Transport Equations	59
3.7	Generalized Transport Systems	61
3.8	Maxwell Equations	62
3.9	Specific References	63
3.10	Problems	63
Chapter 4	Collisions	66
4.1	Simple Collision Parameters	67
4.2	Binary Elastic Collisions	68
4.3	Collision Cross Sections	74
4.4	Transfer Collision Integrals	78
4.5	Maxwell Molecule Collisions	82
4.6	Collision Terms for Maxwellian Velocity Distributions	85
4.7	Collision Terms for 13-Moment Velocity Distributions	90
4.8	Momentum Transfer Collision Frequencies	94
4.9	Specific References	100
4.10	Problems	101
Chapter 5	Simplified Transport Equations	104
5.1	Basic Transport Properties	105
5.2	The 5-Moment Approximation	109
5.3	Transport in a Weakly Ionized Plasma	110
5.4	Transport in Partially and Fully Ionized Plasmas	116
5.5	Major Ion Diffusion	117
5.6	Polarization Electrostatic Field	118
5.7	Minor Ion Diffusion	120
5.8	Supersonic Ion Outflow	123
5.9	Time-Dependent Plasma Expansion	125
5.10	Diffusion Across B	127
5.11	Electrical Conductivities	129
5.12	Electron Stress and Heat Flow	132
5.13	Ion Stress and Heat Flow	137
5.14	Higher-Order Diffusion Processes	138
5.15	Summary of Appropriate Use of Transport Equations	142
5.16	Specific References	144

5.17	General References	145
5.18	Problems	145
Chapter 6	Wave Phenomena	148
6.1	General Wave Properties	148
6.2	Plasma Dynamics	153
6.3	Electron Plasma Waves	157
6.4	Ion-Acoustic Waves	158
6.5	Upper Hybrid Oscillations	160
6.6	Lower Hybrid Oscillations	162
6.7	Ion-Cyclotron Waves	163
6.8	Electromagnetic Waves in a Plasma	164
6.9	Ordinary and Extraordinary Waves	167
6.10	L and R Waves	170
6.11	Alfvén and Magnetosonic Waves	172
6.12	Effect of Collisions	173
6.13	Two-Stream Instability	174
6.14	Shock Waves	177
6.15	Double Layers	182
6.16	Summary of Important Formulas	187
6.17	Specific References	189
6.18	General References	190
6.19	Problems	190
Chapter 7	Magnetohydrodynamic Formulation	192
7.1	General MHD Equations	193
7.2	Generalized Ohm's Law	197
7.3	Simplified MHD Equations	198
7.4	Pressure Balance	199
7.5	Magnetic Diffusion	201
7.6	Spiral Magnetic Field	202
7.7	Double-Adiabatic Energy Equations	204
7.8	Alfvén and Magnetosonic Waves	206
7.9	Shocks and Discontinuities	209
7.10	Specific References	213

7.11	General References	213
7.12	Problems	214
Chapter 8 Chemical Processes 216		
8.1	Chemical Kinetics	216
8.2	Reaction Rates	220
8.3	Charge Exchange Processes	223
8.4	Recombination Reactions	227
8.5	Negative Ion Chemistry	229
8.6	Excited State Ion Chemistry	231
8.7	Optical Emissions; Airglow and Aurora	232
8.8	Specific References	234
8.9	General References	235
8.10	Problems	235
Chapter 9 Ionization and Energy Exchange Processes 237		
9.1	Absorption of Solar Radiation	237
9.2	Solar EUV Intensities and Absorption Cross Sections	241
9.3	Photoionization	242
9.4	Superthermal Electron Transport	246
9.5	Superthermal Ion and Neutral Particle Transport	251
9.6	Electron and Ion Heating Rates	254
9.7	Electron and Ion Cooling Rates	258
9.8	Specific References	265
9.9	General References	267
9.10	Problems	267
Chapter 10 Neutral Atmospheres 269		
10.1	Rotating Atmospheres	270
10.2	Euler Equations	271
10.3	Navier–Stokes Equations	272
10.4	Atmospheric Waves	274
10.5	Gravity Waves	275
10.6	Tides	279
10.7	Density Structure and Controlling Processes	283

- 10.8 Escape of Terrestrial Hydrogen 290
- 10.9 Energetics and Thermal Structure of the Earth's Thermosphere 293
- 10.10 Exosphere 299
- 10.11 Hot Atoms 304
- 10.12 Specific References 306
- 10.13 General References 308
- 10.14 Problems 308

Chapter 11 **The Terrestrial Ionosphere at Middle and Low Latitudes** 312

- 11.1 Dipole Magnetic Field 314
- 11.2 Geomagnetic Field 318
- 11.3 Geomagnetic Variations 320
- 11.4 Ionospheric Layers 323
- 11.5 Topside Ionosphere and Plasmasphere 331
- 11.6 Plasma Thermal Structure 336
- 11.7 Diurnal Variation at Mid-Latitudes 342
- 11.8 Seasonal Variation at Mid-Latitudes 344
- 11.9 Solar Cycle Variation at Mid-Latitudes 345
- 11.10 Plasma Transport in a Dipole Magnetic Field 346
- 11.11 Equatorial *F* Region 347
- 11.12 Equatorial Spread *F* and Bubbles 350
- 11.13 Sporadic *E* and Intermediate Layers 355
- 11.14 Tides and Gravity Waves 357
- 11.15 Ionospheric Storms 360
- 11.16 Specific References 361
- 11.17 General References 364
- 11.18 Problems 364

Chapter 12 **The Terrestrial Ionosphere at High Latitudes** 366

- 12.1 Convection Electric Fields 367
- 12.2 Convection Models 374
- 12.3 Effects of Convection 378
- 12.4 Particle Precipitation 386
- 12.5 Current Systems 391

12.6	Large-Scale Ionospheric Features	393
12.7	Propagating Plasma Patches	397
12.8	Boundary and Auroral Blobs	399
12.9	Sun-Aligned Arcs	401
12.10	Geomagnetic Storms	402
12.11	Substorms	404
12.12	Polar Wind	406
12.13	Energetic Ion Outflow	421
12.14	Specific References	426
12.15	General References	430
12.16	Problems	431
Chapter 13	Planetary Ionospheres	433
13.1	Mercury	433
13.2	Venus	433
13.3	Mars	443
13.4	Jupiter	445
13.5	Saturn, Uranus, Neptune, and Pluto	447
13.6	Satellites and Comets	451
13.7	Specific References	458
13.8	General References	461
13.9	Problems	462
Chapter 14	Ionospheric Measurement Techniques	464
14.1	Spacecraft Potential	464
14.2	Langmuir Probes	466
14.3	Retarding Potential Analyzers	469
14.4	Thermal Ion Mass Spectrometers	472
14.5	Radio Reflection	475
14.6	Radio Occultation	477
14.7	Incoherent (Thomson) Radar Backscatter	480
14.8	Specific References	485
14.9	General References	488

Appendices 489**A Physical Constants and Conversions 489****B Vector Relations and Operations 491**

B.1 Vector Relations 491

B.2 Vector Operators 492

C Integrals and Transformations 494

C.1 Integral Relations 494

C.2 Important Integrals 495

C.3 Integral Transformations 495

D Functions and Series Expansions 497

D.1 Important Functions 497

D.2 Series Expansions for Small Arguments 498

E Systems of Units 499

Table E.1 Widely used formulas 500

F Maxwell Transfer Equations 501**G Collision Models 506**

G.1 Boltzmann Collision Integral 506

G.2 Fokker–Planck Collision Term 510

G.3 Charge Exchange Collision Integral 510

G.4 Krook Collision Models 511

G.5 Specific References 512

H Maxwell Velocity Distribution 513

Specific Reference 517

I	Semilinear Expressions for Transport Coefficients	518
I.1	Diffusion Coefficients and Thermal Conductivities	518
I.2	Fully Ionized Plasma	519
I.3	Partially Ionized Plasma	520
I.4	Specific References	520
J	Solar Fluxes and Relevant Cross Sections	521
Table J.1	Parameters for the EUVAC Solar Flux Model	522
Table J.2	Photoabsorption and photoionization cross sections	523
	Specific References	531
K	Atmospheric Models	532
K.1	Introduction	532
Table K.1	VIRA model of composition, temperature, and density (Noon)	533
Table K.2	VIRA model of composition, temperature, and density (Midnight)	534
Table K.3	MSIS model of terrestrial neutral parameters (Noon, Winter)	535
Table K.4	MSIS model of terrestrial neutral parameters (Midnight, Winter)	536
Table K.5	MSIS model of terrestrial neutral parameters (Noon, Summer)	537
Table K.6	MSIS model of terrestrial neutral parameters (Midnight, Summer)	538
K.2	Specific References	538
L	Scalars, Vectors, Dyadics and Tensors	539
	Index	545

Chapter 1

Introduction

1.1 Background and Purpose

The ionosphere is considered to be that region of an atmosphere where significant numbers of free thermal (<1 eV) electrons and ions are present. All bodies in our solar system that have a surrounding neutral-gas envelope, due either to gravitational attraction (e.g., planets) or some other process such as sublimation (e.g., comets), have an ionosphere. Currently, ionospheres have been observed around all but two of the planets, some moons, and comets. The free electrons and ions are produced via ionization of the neutral particles both by extreme ultraviolet radiation from the Sun and by collisions with energetic particles that penetrate the atmosphere. Once formed, the charged particles are affected by a myriad of processes, including chemical reactions, diffusion, wave disturbances, plasma instabilities, and transport due to electric and magnetic fields. Hence, an understanding of ionospheric phenomena requires a knowledge of several disciplines, including plasma physics, chemical kinetics, atomic theory, and fluid mechanics. In this book, we have attempted to bridge the gaps among these disciplines and provide a comprehensive description of the physical and chemical processes that affect the behavior of ionospheres.

A brief history of ionospheric research is given later in this introductory chapter. An overview of the space environment, including the Sun, planets, moons, and comets, is presented in Chapter 2. This not only gives the reader a quick look at the overall picture, but also provides the motivation for the presentation of the material that follows. Next, in Chapter 3, the general transport equations for mass, momentum, and energy conservation are derived from first principles so that the reader can clearly see where these equations come from. This is followed by a derivation of the collision terms that appear in the transport equations, including those relevant to resonant charge

exchange, nonresonant ion-neutral and electron-neutral interactions, and Coulomb collisions (Chapter 4). These general collision terms and transport equations are complicated and in many situations it is possible to use simpler sets of transport equations. Therefore, in Chapter 5, several simplified systems of transport equations are derived, including the Euler, Navier–Stokes, diffusion, and thermal conduction equations. This is followed by a discussion of the wave modes, plasma instabilities, and shocks that can occur in the ionospheres (Chapter 6). In Chapter 7, the magnetohydrodynamic (MHD) equations are derived and then used to describe MHD waves, shocks, and pressure balance.

In Chapter 8, chemical kinetics and a variety of reactions relevant to the ionospheres are discussed and presented, including those involving metastable species and negative ions. Optical emissions are also briefly discussed in this chapter. The relevant ionization and energy exchange processes are detailed in Chapter 9, including those pertaining to both photons and particles. The chapter concludes with a summary of the heating and cooling expressions that are needed for practical applications. Chapter 10 is devoted to a discussion of neutral atmospheres. The Euler and Navier–Stokes equations for neutral gases are presented at the beginning of the chapter, and this is followed by a discussion of atmospheric waves and tides. The rest of Chapter 10 deals with atmospheric structure, escape fluxes, the exosphere, and hot atoms. In Chapters 11 and 12, the general material given in the previous chapters is applied to elucidate the unique characteristics associated with the terrestrial ionosphere at low, middle, and high latitudes. Although much of this material is still of a fundamental nature, an overview of what has been accomplished to date is also provided. Chapter 13 summarizes what is currently known about all of the other ionospheres in the solar system. The most commonly used experimental techniques for measuring ionospheric densities, temperatures, and drifts are briefly described in Chapter 14. Finally, several Appendices are included that contain physical constants, mathematical formulas, some important derivations, and useful tables.

This book is the outgrowth of two decades of numerous joint research endeavors and publications by the authors. Some of the material was used in courses taught by the authors at Utah State University and at the University of Michigan. This book should be useful to graduate students, postdoctoral fellows, and established scientists who want to fill gaps in their knowledge. It also serves as a reference book for obtaining important equations and formulas. A subset of the material can be used for a graduate level course about the upper atmosphere and ionosphere and/or plasma physics. At the University of Michigan a one semester graduate course on the ionosphere and upper atmosphere has been based on Chapters 2, 3, parts of 5, 8, 9, 10, most of 11 and 12, and 13 and 14. At Utah State University, a one semester course on plasma physics has been based on Chapters 3–7, and a course on aeronomy has been based on Chapters 2, 3, 5, 8–12. To facilitate the use of this book as a text, problems are provided at the end of most of the chapters.

Several people were helpful in the preparation of this book, and we wish to acknowledge them here. The help came in a variety of forms (e.g., providing some

unpublished material, reading and/or proofing part of the manuscript, etc.), and it certainly improved the book. AFN would especially like to thank (in alphabetical order) J. R. Barker, T. E. Cravens, J. L. Fox, B. E. Gilchrist, T. I. Gombosi, J. W. Holt, A. J. Kliore, M. W. Liemohn, and H. Rishbeth. RWS would like to thank Melanie Oldroyd for typing a preliminary form of some of the chapters. Both of us would like to thank Shawna Johnson for drawing some of the figures, for digitizing figures, and for overseeing the production of the book. Both of us would also like to thank Elizabeth Wood for preparing the manuscript in L^AT_EX. Some of the material in the book comes from lecture notes collected over many years and thus may contain material without appropriate references to their sources, which we have forgotten. This is inadvertent and we apologize to such authors. Also, in order to keep the bibliographies from becoming unrealistically long, we limited our referencing to only those papers from which figures were taken, to either the latest or original reference for the material discussed, and to review papers. Hence, we omitted many deserving, appropriate, and relevant references. We hope that the readers and scientists working in the field will understand and appreciate our dilemma.

The units used in the book are a mixture of MKSA and Gaussian-cgs because of the corresponding usage by practitioners in the field. All equations and formulas throughout the book are in MKSA units, but some tables and numbers are given in Gaussian-cgs units when this is the common practice. The conversion from one system to the other is briefly discussed in Appendix E.

1.2 History of Ionospheric Research

The earliest exposure of mankind to a phenomenon originating in the upper atmosphere is the visual aurora. The visual displays of colored light appear in the form of arcs, bands, patches, blankets, and rays, and often the features move rapidly across the night sky. It has been suggested that the earliest records of the aurora can be traced to Stone Age man.¹ References to the aurora appear in the Old Testament, in writings of Greek philosophers including Aristotle's *Meteorologica*, and possibly in ancient Chinese works before 2000 B.C. In most of these early writings, the auroral displays were interpreted to be manifestations of God. The name *aurora borealis* (northern dawn) appears to have been coined by Galileo at some time prior to 1621.¹ The first recorded observation of the southern hemispheric aurora (*aurora australis*) was by Cook in 1773.

A serious scientific study of auroras began at about 1500 A.D.¹ However, the early theories put forth by noted scientists were completely wrong. Edmund Halley, who predicted the reappearance of what is now known as Halley's comet, suggested that the auroras were "watery vapors, which are rarefied and sublimed by subterraneous fire, [and] might carry along with them sulphureous vapors sufficient to produce this luminous appearance in the atmosphere." In 1746, the Swiss mathematician Leonard Euler suggested that "the aurora was particles from the Earth's own atmosphere driven

beyond its limits by the impulse of the sun's light and ascending to a height of several thousand miles. Near the poles, these particles would not be dispersed by the Earth's rotation."² Benjamin Franklin, who was a respected scientist in his time, thought that the aurora was related to atmospheric circulation patterns.³ Basically, Franklin argued that the atmosphere in the polar regions must be heavier and lower than in the equatorial region because of the smaller centrifugal force, and therefore, the vacuum-atmosphere interface must be lower in the polar regions. He then further argued that the electricity brought into the polar region by clouds would not be able to penetrate the ice, and hence, would break through the low atmosphere and run along the vacuum toward the equator. The electricity would be most visible at high latitudes, where it is dense, and much less visible at lower latitudes, where it diverges. Franklin claimed such an effect would "give all the appearances of an Aurora Borealis."¹

Numerous other theories of the aurora have been proposed over the last 150 years, including reflected sunlight from ice particles, reflected sunlight from clouds, sulfurous vapors, combustion of inflammable air, luminous magnetic particles, meteoric dust ignited by friction with the atmosphere, cosmic dust, currents generated by compressed cosmic ether, thunderstorms, electric discharges between the Earth's magnetic poles, and electric discharges between fine ice needles. A comprehensive and fascinating account of the aurora in science, history, and the arts is given in Reference 1, and additional theories are presented there.

Although early auroral theories did not fare very well, observations made during the latter half of the 1700s and throughout the 1800s elucidated many important auroral characteristics. In 1790, the English scientist Cavendish used triangulation and estimated the height of auroras at between 52 and 71 miles.⁴ In 1852, the relationship among geomagnetic disturbances, auroral displays, and sunspots was clearly established; the frequency and amplitude of these features varied with the same 11-year periodicity.^{5,6} In 1860, Elias Loomis drew the first diagram of the region where auroras are most frequently observed and noted that the narrow ring is not centered on the geographic pole, but that its oval form resembles lines of equal magnetic dip, thereby establishing the relationship between the aurora and the geomagnetic field. In 1867, the Swedish physicist Angström made the first measurements of the auroral spectrum.⁷ However, a significant breakthrough in auroral physics was not achieved until the end of the nineteenth century, when cathode rays were discovered and identified as electrons by the British physicist J. J. Thomson. Subsequently, the Norwegian physicist Kristian Birkeland proposed that the aurora was caused by a beam of electrons emitted by the Sun. Those electrons reaching the Earth would be affected by the Earth's magnetic field and guided to the high-latitude regions to create the aurora.

Until the discovery of sunspots by Galileo in 1610, the Sun was generally thought to be a quiet, featureless object. Galileo not only discovered the dark spots but also noted their westward movement, which was the first indication that the Sun rotates. In subsequent observations, it was quickly established that the number of sunspots varies with time. It was not until more than two centuries later, however, that an amateur astronomer in Germany, Heinrich Schwabe, noted an apparent 10-year periodicity in his 17 years of sunspot observations.⁸ Shortly after Schwabe's discovery, professional

astronomers set out to determine whether or not the cycle was real. The leader of this effort was Rudolf Wolf of the Zürich observatory. Wolf conducted an extensive search of past data and was able to establish that the number of sunspots varied with an 11-year cycle that had been present since at least 1700.⁹ In 1890, Maunder called attention to the 70-year period from 1645 to 1715, when almost no sunspots were observed.¹⁰ This period, which is known as the Maunder Minimum, raises the question whether the sunspot cycle is a universal feature or just a recent phenomenon.

As defined at the beginning of this section, the terrestrial ionosphere begins at an altitude of about 60 km and extends beyond 3000 km, with the peak electron concentration occurring at approximately 300 km. The first suggestion of the existence of what is now called the ionosphere can be traced to the 1800s. Carl Gauss and Balfour Stewart hypothesized the existence of electric currents in the atmosphere to explain the observed variations of the magnetic field at the surface of the Earth. Gauss argued:¹¹

It may indeed be doubted whether the seat of the proximate causes of the regular and irregular changes which are hourly taking place in this [terrestrial magnetic] force, may not be regarded as external in reference to the Earth . . . But the atmosphere is no conductor of such [galvanic] currents, neither is vacant space. But our ignorance gives us no right absolutely to deny the possibility of such currents; we are forbidden to do so by the enigmatic phenomena of the Aurora Borealis, in which there is every appearance that electricity in motion performs a principal part.

It had been well established that there was a direct correlation between the solar cycle and magnetic disturbances on the Earth. To account for this strong correlation, Stewart speculated that electrical currents must flow in the Earth's upper atmosphere, and that the Sun's action is responsible for turning air into a conducting medium.¹² It was also concluded that the conductivity of the upper atmosphere is higher at sunspot maximum than at sunspot minimum. This view, however, was not widely accepted and strong counterarguments were presented in 1892 by Lord Kelvin.

The existence of the ionosphere was clearly established in 1901 when G. Marconi successfully transmitted radio signals across the Atlantic. This experiment indicated that radio waves were deflected around the Earth's surface to a much greater extent than could be attributed to diffraction. The following year, A. E. Kennelly and O. Heaviside suggested that free electrical charges in the upper atmosphere could reflect radio waves.¹³ That same year, the first *physical* theory of the ionosphere was proposed.¹⁴

The observed effect, which if confirmed is very interesting, seems to me to be due to the conductivity . . . of air, under the influence of ultra-violet solar radiation. No doubt electrons must be given off from matter . . . in the solar beams; and the presence of these will convert the atmosphere into a feeble conductor.

In 1903, J. E. Taylor independently suggested that solar ultraviolet radiation was the source of electrical charges, which implied solar control of radio propagation.¹⁵ The first rough measurements of the height of the reflecting layer were made by Lee de Forest and L. F. Fuller at the Federal Telegraph Company in San Francisco from 1912 to 1914. The reflecting layer's height was deduced using a transmitter-receiver

spacing of approximately 500 km, which was determined by the circuits of the Federal Telegraph Company.¹⁶ However, the de Forest–Fuller results were not well known, and generally accepted measurements of the height of the reflecting layer were made in 1924 by Breit and Tuve¹⁷ and by Appleton and Barnett.¹⁸ The Breit–Tuve experiments involved a “pulse sounding” technique, which is still in use today, while Appleton and Barnett used “frequency change” experiments, which demonstrated the existence of downcoming waves by an interference technique. These experiments led to a considerable amount of theoretical work, and in 1926 the name “ionosphere” was proposed by R. A. Watson-Watt in a letter to the United Kingdom Radio Research Board, but it did not appear in the literature until three years later.¹⁹ Radio soundings of the ionosphere initially seemed to indicate that the ionosphere consisted of distinct layers; we now know that this is generally not the case and we refer to different regions. These regions are called the *D*, *E*, and *F* regions. The names of these regions originated with Appleton, who stated that in his early work he wrote *E* for the reflected electric field from the first layer that he recognized. Later, when he recognized a second layer at higher altitudes, he wrote *F* for the reflected field. Subsequently, he conjectured that there may be another layer at lower altitudes so he decided to name the first two layers *E* and *F* and the possible lower one *D*, thus allowing the alphabetical designation of other undiscovered layers.²⁰

The rocket technology available at the end of World War II was used by scientists to study the upper atmosphere and ionosphere, paving the way for space exploration via satellites. The first rocket-borne scientific payload, which carried instrumentation to make measurements directly in the upper atmosphere and ionosphere, was launched in 1946 on a *V-2* from White Sands. The University of Michigan payload consisted of a Langmuir probe and a thermionic pressure gauge; although the *V-2* failed during this flight it marked the beginning of direct exploration of the ionosphere. The first book devoted to the ionosphere was published in 1952 by Rawer.²¹

The rocket technology, coupled with a major advance in ground-based instrumentation, led scientists to realize that a dramatic increase in our knowledge of the terrestrial environment was possible. To take advantage of these new capabilities, the International Geophysical Year (IGY), 1957–1958, was organized.^{22,23} This cooperative effort was to begin with the next maximum of the solar cycle. As part of the IGY, scientists proposed to launch artificial satellites, and eventually *Sputnik 1* was launched on October 4, 1957.

Many consider the launch of *Sputnik 1* the beginning of the Space Age, but to some degree it started much earlier. Rockets have been with us ever since the ancient Chinese used them for fireworks. Later variations of “rockets” were used, basically for military purposes, to send payloads from one location to impact at another. Newton developed the scientific basis to describe how an object could be placed in orbit around the Earth, and visionaries like Jules Verne and H. G. Wells dreamt such thoughts.

The modern era of rocket propulsion began in Russia in the 1880s, where Konstantin Tsiolkovsky worked out the fundamental laws of rocket propulsion and published his work proving the feasibility of achieving orbital velocities by rockets at the turn of the century. He had earlier described the phenomenon of weightlessness in space, predicted

Earth satellites, and suggested the use of liquid hydrogen and oxygen as propellants. Robert H. Goddard, a high school physics teacher in Massachusetts, was not aware of Tsiolkovsky's work and independently began studying rocket propulsion after World War I. On March 16, 1926, he launched the first liquid fuel rocket, which burned for only 2.5 seconds and landed a couple hundred feet away from the "launch site." He continued to work, supported by the Guggenheim Foundation, in seclusion from the press, which ridiculed him. The third rocket pioneer was Hermann Oberth of Germany, also a school teacher. His work gained a great deal of attention and support and eventually led to the development of the V-2 rocket (the first operational liquid fuel rocket).

After World War II, part of the German team responsible for the development of the V-2, including Wernher von Braun, came to the United States, while others went to the Soviet Union. Some of the captured V-2 rockets that were brought to the United States were used to carry scientific payloads; these flights started in May 1946 from White Sands, New Mexico. A year later, the first Soviet V-2 was launched from Kapustin-Yar. The limited supplies of V-2s and the estimated large expense of reproducing them led to the development of new sounding rockets for scientific research. The first one of these was the liquid fuel *Aerobee*; other rockets, many of them having a military heritage, followed later. The ascent of the Cold War spurred the development of Intercontinental Ballistic Missiles (ICBMs), but much of this effort was highly secret. Reading the history of repetitive studies, interservice jealousies, politicking, backbiting, and bickering provides a fascinating view of this secret world of the 1950s.^{24,25}

The first U.S. study regarding the feasibility of artificial satellites can be traced to 1945 when a Navy committee concluded that they were possible, but nothing developed at that time. After a failed Navy–Air Force collaborative effort, the Air Force conducted an independent study and concluded that the United States could launch a 500 pound satellite by 1951. Again, this suggestion was not pursued. However, with pressure from the American Rocket Society and the scientists involved in planning for the IGY, serious consideration was at last given to the launch of a small satellite for scientific purposes. Specifically, on July 29, 1955, a White House announcement indicated that the United States would launch "small unmanned Earth-circling satellites as part of the U.S. participation in the IGY." Two days later, the Soviet Union announced it would also launch artificial satellites as part of the IGY in the late summer or early autumn of 1957. However, this announcement was basically ignored by the U.S. press and public, possibly because it was believed that the Russians did not have the required technology. In the United States, all three military services proposed to launch the first satellite, which was to be placed in orbit during the 1957–1958 time period. The Air Force proposed to use the *Atlas* ICBM, the Army proposed to use the *Jupiter C* Intermediate Range Ballistic Missile (IRBM), and the Navy proposed to develop a new rocket that did not have a military heritage (the *Vanguard*). The *Vanguard* Project was chosen primarily because it would not interfere with the existing military missile programs and because it seemed more appropriate to use a nonmilitary missile for a scientific mission. Despite not being selected, the Army's design of the *Jupiter C* IRBM contained a fourth stage, which appeared to have no specific military function.

In September 1956, when the *Jupiter C* was ready to be launched, the Pentagon was so concerned that the Army might take “the glory” away from the Navy’s Vanguard Program that von Braun was personally ordered to make sure that the fourth stage was not live. The launch was successful, and with a live fourth stage, the *Jupiter C* could have placed a satellite in orbit.

On October 4, 1957, the Soviet Union launched *Sputnik (“Traveler/Companion”) 1*, which was an 83 kg satellite. *Sputnik 2*, a 507-kg satellite followed on November 3. This created a tremendous public and political reaction in the United States.

Vanguard was still given a first chance, but the launch attempt on December 6, 1957, was a televised public failure (the second launch attempt on February 5, 1958, was also a failure). In the meantime the Army was given the green light to proceed with a *Jupiter C* launch, and an 8-kg satellite named *Explorer I* was successfully placed in orbit on January 31, 1958. *Explorer I* carried a small Geiger counter supplied by James Van Allen of the University of Iowa. The instrument was supposed to record the presence of cosmic rays, which are very fast particles from deep space; but surprisingly the instrument showed no response when the satellite was at high altitudes. There seemed to be no logical explanation, but a second instrument flown two months later confirmed the result. A graduate student working with Van Allen solved the problem. He suggested that the satellite encountered a region of very intense energetic particles, which saturated the Geiger tube and caused the counting circuits to read zero. Thus, the Van Allen radiation belts were discovered.²⁶

The large international cooperative efforts, the vast amount of geophysical data collected, and the launch of artificial satellites, which began during the IGY, led to the birth of solar-terrestrial physics. The subsequent major infusion of money into this area by several countries led to a rapid advance in our knowledge of the Earth’s environment. In the early phase of these explorations, every measurement yielded new and exciting results. A phase has now been reached where detailed measurements are available and theoretical models are generally able to explain and reproduce the observed large-scale features of the terrestrial ionosphere. This does not imply that a complete understanding has been achieved and there is nothing more to learn. On the contrary, the time has been reached when the problems that need further study can be clearly defined and then attacked in a systematic manner.

1.3 Specific References

1. Eather, R. H., *Majestic Lights*, American Geophysical Union, 1980.
2. Euler, L., Recherches physiques sur la cause des queues comètes de la lumière boreale et de la lumière zodiacale, *Hist. Acad. Roy. Sci. Belles Lett. Berlin* **2**, 117, 1746.
3. Franklin, B., Political, Miscellaneous and Philosophical Pieces, ed. by Vaughan, 504, Johnson, 1779.
4. Cavendish, H., On the height of the luminous arch which was seen on February 23, 1784, *Phil. Trans. Roy. Soc.* **80**, 101, 1790.

5. Sabine E., On periodical laws discoverable in the mean effects of the larger magnetic disturbances, *Phil. Trans. Roy. Soc.*, **142**, 103, 1852.
6. Wolf, R., *Acad. Sci.*, **35**, 364, 1852.
7. Angström, A., Spectrum des Nordlichts, *Ann. Phys.* **137**, 161, 1869.
8. Schwabe, S. H., Die Sonne, *Astron. Nachr.*, **20**, 280, 1843.
9. Wolf, R., Neue Untersuchungen über die Periode der Sonnenflecken, *Astron. Mitt. Zürich*, No. 1, 8, 1856.
10. Maunder, E. W., The prolonged sunspot minimum, 1645–1715, *MNRAS*, **50**, 251, 1890.
11. Gauss, C. F., General theory of terrestrial magnetism, English translation in *Scientific Memoirs*, (ed. R. Taylor), Vol. 2, 184, London, 1841.
12. Stewart, B., Aurora Borealis, in *Encyclopaedia Britannica*, 9th Ed., 36, 1882.
13. Ratcliffe, J. A., The ionosphere and the engineer, *Proc. Inst. Elec. Eng. (London)*, **114**, 1, 1967.
14. Lodge, O., Mr. Marconi's results in day and night wireless telegraphy, *Nature*, **66**, 222, 1902.
15. Taylor, J. E., Characteristics of electric earth-current disturbances, and their origin, *Proc. Phys. Soc. (London)*, **LXXI**, 225, 1903.
16. Villard, O. G., The ionospheric sounder and its place in the history of radio science, *Radio Sci.*, **11**, 847, 1976.
17. Breit, G., and M. A. Tuve, A radio method of estimating the height of the conducting layer, *Nature*, **116**, 357, 1925.
18. Appleton, E. V., and M. A. F. Barnett, Local reflection of wireless waves from the upper atmosphere, *Nature*, **115**, 333, 1925.
19. Watson-Watt, R. A., Weather and wireless, *Q. J. Roy. Meteorol. Soc.*, **55**, 273, 1929.
20. Silberstein, R., The origin of the current nomenclature for the ionospheric layers, *J. Atmos. Terr. Phys.*, **13**, 382, 1959.
21. Rawer, K., *Die Ionosphere*, P. Noordhoff Ltd., Groningen, 1952.
22. Van Allen, J. A., Genesis of the International Geophysical Year, in *History of Geophysics*, ed. by C. S. Gillmor, American Geophysical Union, **4**, 49, 1984.
23. Nicolet, M., Historical aspects of the IGY, in *History of Geophysics*, ed. by C. S. Gillmor, American Geophysical Union, **44**, 44, 1984.
24. Emme, E. M., *Aeronautics and Astronautics, 1915–1960*, U.S. Government Printing Office, Washington, D.C., 1961.
25. Thomas, S., *Men in Space*, Vol. 2, 218, Chilton Press, Philadelphia, 1961.
26. Van Allen, J. A., Radiation belts around the earth, *Sci. Amer.*, **200**, 39, 1959.

1.4 General References

- Eather, R. H., *Majestic Lights*, American Geophysical Union, Washington, D.C., 1980.
- Hess, W. N., *The Radiation Belt and Magnetosphere*, Blaisdell Publishing Co., Waltham, MA, 1968.

Newell, H. E., *Beyond the Atmosphere*, NASA SP-4211, Washington, D.C., 1980.

Rishbeth, H., H. Kohl, and L. W. Barclay, A history of ionospheric research and radio communications, in *Modern Ionospheric Science*, European Geophysical Society, Katlenburg-Lindau, FRG, 1996.

Van Allen, J. A., *Origins of Magnetospheric Physics*, Smithsonian Institution Press, Washington, D.C., 1983.