Lightning-excited, Quasi-electrostatic Effects in the Lower Ionosphere

by

James Ernstmeyer

S.B. Physics, Massachusetts Institute of Technology, 1984
S.M. Physics, Massachusetts Institute of Technology, 1984

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Science

at the

Massachusetts Institute of Technology

June 1998

© 1998 Massachusetts Institute of Technology. All rights reserved

Signature of Author: ____________________________

Department of Electrical Engineering and Computer Science

April 23, 1998

Certified by: ____________________________

Tom T. S. Chang
Director, Center for Theoretical Geo/Cosmo Plasma Physics
Thesis Co-Supervisor

Certified by: ____________________________

Jin Au Kong
Professor of Electrical Engineering
Thesis Co-Supervisor

Accepted by: ____________________________

Arthur C. Smith
Chairman, Department Committee on Graduate Students

JUL 23 1998

ARCHIVE

LIBRARIES
Lightning-excited, Quasi-electrostatic Effects in the Lower Ionosphere

by

James Ernstmeyer

Submitted to the Department of Electrical Engineering and Computer Science on April 23, 1998
in partial fulfillment of the requirements for the degree of
Doctor of Science in Electromagnetics

Abstract

A study is presented of the quasi-electrostatic effects on the mesosphere and lower ionosphere due to lightning discharges. A comprehensive, three-dimensional theory and computational model is developed to explain the structure and occurrence of heated electron populations in the lower ionosphere after a lightning discharge. The computational model is applied under conditions that are associated with the formation of the high altitude luminous phenomena known as red sprites, and the results are used to elucidate the underlying mechanisms. A one-dimensional, nonlinear analytical theory is developed to explain propagating density discontinuities in the mesospheric-ionospheric medium.

A comprehensive, multiple-fluid transport model is developed. Atmospheric and ionospheric conditions are accurately represented in a cylindrically symmetric simulation volume that extends from the ground to 90 km altitude, and to a radius of 64 km. Cloud electrification and lightning discharge comprise the stimulus to the system. Self-consistently calculated outputs include the electric field vector, charge density, current density, electron density, negative ion density, small positive ion density, positive cluster ion density, and electron temperature. The electrostatic field is obtained by combining a low-altitude numerical solution to Poisson's Equation with a more accurate quasi-neutral analytical solution at high altitudes. Electron flux is computed from a steady state solution to the momentum equation. Ion fluxes are found by numerical solution of the ion momentum equations. The electron temperature is obtained from a steady state solution to the electron energy transport equation, including the effects of elastic and inelastic energy loss processes.

Detailed collisional and chemical effects models are developed. A hybrid, fluid/kinetic technique is used to compute the collision, excitation, and ionization rates for use in the fluid formulation. The electron distribution function is obtained from a kinetic theory solution in the presence of a static electric field. This distribution is employed in conjunction with measured cross sections to obtain the required collisional process rates. Chemical effects are calculated from a system of equations that incorporate impact ionization, cosmic ray ionization, three-body attachment of electrons, dissociative attachment of electrons to oxygen, hydration, dehydration, and recombination of all species. The equilibrium densities and dynamic process rates obtained in this manner are valid at all altitudes in the simulation volume.

The numerical solution of the three-dimensional model is presented. The flux-corrected transport method is extended for use in a cylindrical geometry, and is validated for application in a volume bounded by the earth and the ionosphere. Positive cloud-to-ground lightning is shown to produce high altitude electron heating to 0.7 eV, consistent with the formation of sprites. High altitude charge separation is found to be associated with positive, but not negative, cloud-to-ground
lightning discharges, a fact which helps to explain the strong correlation between positive discharges and sprites. Redistribution of electron and ion populations in the 20 to 40 km altitude region is shown to persist for tens of seconds after a lightning discharge. Two new mesospheric phenomena are predicted by the simulation. The first is a band of depressed electron temperature that moves rapidly downward from 70 to 25 km during the first few milliseconds after the onset of a lightning discharge. The second is a discontinuity in the electron density that propagates downward along the simulation axis for 200 to 300 ms after a discharge. This discontinuity can be understood as a steady-state attachment wave.

A one-dimensional, nonlinear fluid model of attachment wave propagation is developed. Three species, namely electrons and positive and negative ions, are included in the model, and the electron temperature varies self-consistently according to a simplified energy transport equation. The electrostatic field is determined by the Poisson Equation. Analysis of the system near its equilibrium points is used to establish conditions for propagation, depending upon the sign of the asymptotic low altitude electric field. The model suggests that a stationary attachment wave front is maintained under a balance between drift and advection of the charged populations.

Thesis Co-Supervisor: Tom T. S. Chang
Title: Director, Center for Theoretical Geo/Cosmo Plasma Physics

Thesis Co-Supervisor: Jin Au Kong
Title: Professor of Electrical Engineering
Acknowledgments

I owe the deepest gratitude to Dr. Tom Chang, whose faith in me over a period of many years has made this work possible. His kindness and patience will always inspire me.

I wish to thank Prof. Jin Kong for his encouragement and his support of a cooperative effort with Dr. Chang. I am thankful to Prof. Markus Zahn, who graciously volunteered to serve on my thesis committee, and to Prof Qing Hu, who provided me a valuable experience in solid state physics.

I am especially thankful to Prof. Nikos Gatsonis of the Worcester Polytechnic Institute for generously offering his flux-corrected transport code, which helped me to overcome a major hurdle. Heartfelt thanks go also to Dr. Dwight Decker of the Air Force Research Laboratory, for providing me with collisional cross section code. I appreciate the many encouraging and helpful discussions I had with Drs. John Retterer, Jack Jasperse, Dick Picard, and Jeremy Winick of the Air Force Research Laboratory, and Dr. Peter Wintersteiner of ARCON Corporation.

I owe special thanks to Delia Donatelli, who invited me to join the Rome Air Development Center, thereby setting the stage for the many triumphs and tests of will that followed. I extend sincere appreciation to Stuart Matthews, Ray Lund, John Du Long, Ty Jarry, Tim Tanigawa, and Neil Myers, who capably brought the Ducted Wave Experiment to its conclusion in my absence, and to Bert Weijers for his steadfast support.

I was supported by the Sensors Directorate, Air Force Research Laboratory, for the duration of this effort.
## Contents

List of Figures ............................................. 9  
List of Tables ............................................. 11  
List of Symbols ............................................ 12

1 Introduction ............................................... 16
   1.1 Observations ........................................... 18  
       1.1.1 Sprites and Related Phenomena ......................... 18  
       1.1.2 Lightning Phenomenology ............................... 25  
   1.2 Previous Simulation Approaches ......................... 27  
   1.3 Distinction of the Present Approach .................... 33

2 Three-dimensional Simulation Model .......................... 35
   2.1 Overview and Geometry ................................... 35  
   2.2 Modeling the Ambient Medium ............................ 39  
       2.2.1 The Neutral Atmosphere .............................. 39  
       2.2.2 The Charged Constituents ............................ 41  
   2.3 Cloud Charge and Lightning Model ....................... 46  
   2.4 Electric Field Model .................................... 50  
       2.4.1 Fair Weather Electric Field .......................... 51  
       2.4.2 Boundary Conditions .................................. 54  
       2.4.3 High Altitude Electric Field Solution ................ 58  
   2.5 Collisional Effects Model ............................... 65  
       2.5.1 Collision Rates ....................................... 65  
       2.5.2 Electron Distribution Function ....................... 66  
       2.5.3 Collisional Effects Calculated For Maxwellian Versus Druyvestein Electron Distributions ......................... 73  
       2.5.4 Hybrid Fluid-Kinetic Approach ....................... 77  
   2.6 Chemical Effects Model .................................. 79  
       2.6.1 Species, Reactions, and Rates ....................... 80
2.6.2 Rate Equations and Equilibrium Ion Densities .................................................. 86
2.7 Fluid Model .............................................................................................................. 92
  2.7.1 The Continuity Equation .................................................................................... 94
  2.7.2 The Momentum Equation .................................................................................. 96
  2.7.3 The Energy Transport Equation ....................................................................... 98
  2.7.4 Normalization of the Fluid Equations ............................................................... 101
  2.7.5 Steady State Fluid Results ............................................................................... 102

3 Numerical Methods ................................................................................................... 106
  3.1 Overview of the Solution Process ......................................................................... 106
  3.2 The Flux-Corrected Transport Method .................................................................. 112
    3.2.1 Overview of FCT Method ............................................................................. 112
    3.2.2 Extending FCT to Cylindrical Systems ......................................................... 117
    3.2.3 Low-order Algorithm ................................................................................... 118
    3.2.4 High-order Algorithm .................................................................................. 119
    3.2.5 Application of the FCT Method to the Momentum Equation ....................... 120
    3.2.6 Stability and Accuracy of the FCT Method .................................................... 121
  3.3 Finite Differences Solution of the Poisson Equation in Cylindrical Coordinates ... 124
    3.3.1 FISHPAK Subroutine PWSCYL ...................................................................... 124
    3.3.2 Electric Field Calculation ............................................................................. 126
    3.3.3 Discretization Error in the Electric Field Solution .......................................... 126
  3.4 The Runge-Kutta Method ..................................................................................... 129

4 Simulation Results ..................................................................................................... 131
  4.1 Overview of Representative Outputs ..................................................................... 131
    4.1.1 Spatial Variation of Observables .................................................................... 132
    4.1.2 Temporal Evolution of Simulation System .................................................... 142
    4.1.3 Applicability of Simulation Approach ........................................................... 147
  4.2 Comparison to Physical Observations .................................................................... 149
  4.3 Influence of Parameters ....................................................................................... 150
    4.3.1 High Altitude Electron Density ....................................................................... 151
    4.3.2 Magnitude of Cloud Charge Lowered to Ground ............................................ 154
4.3.3 Lightning Stroke Duration .................................................. 156

4.4 Special Phenomena ............................................................... 159
  4.4.1 High Altitude Charge Separation ...................................... 160
  4.4.2 Persistent Effects of a Lightning Discharge ....................... 165
  4.4.3 Downward Moving Temperature Depression ........................ 168
  4.4.4 Downward Propagating Electron Density Discontinuity ........ 171

5 One-dimensional Model of an Attachment Wave .......................... 174
  5.1 Previous Approaches to Ionizing Wave Studies ..................... 174
  5.2 Attachment Wave Model .................................................... 177
    5.2.1 Solutions for Positive Low Altitude Electric Field .......... 187
    5.2.2 Solutions for Negative Low Altitude Electric Field ......... 189

6 Conclusions ............................................................................. 193
  6.1 Contributions to Theory ...................................................... 193
  6.2 Contributions to the Study of Red Sprites and Related Phenomena 196
  6.3 Contributions to Numerical Methods .................................... 198
  6.4 Recommendations for Future Research ................................. 198
    6.4.1 Elaboration of Present Models ...................................... 199
    6.4.2 Additional Investigations ............................................. 200

A Electron-Neutral Collisional Processes ..................................... 202

B Hydration of Small Positive Ions ............................................. 205

C Monopolar Versus Dipolar Cloud Charge Models ........................ 211

Bibliography .............................................................................. 215
**List of Figures**

1.1 Altitude ranges of lightning-excited phenomena in relation to ionospheric and atmospheric layers ................................................................. 22

2.1 Simulation overview and geometry ......................................................................................... 37
2.2 Model of neutral atmosphere employed in this work, obtained from COSPAR International Reference Atmosphere (CIRA) .............................................................................. 40
2.3 Small positive ion density for use as an initial condition in the simulation ................... 44
2.4 Spatial and temporal characteristics of cloud charge and lightning model .................. 47
2.5 Behavior of cloud charge under monopole and dipole models .............................................. 49
2.6 Fair weather electrical properties of the atmosphere .............................................................. 53
2.7 Measures of effectiveness of boundary condition approximations ........................................ 57
2.8 The high altitude electric field solution ..................................................................................... 62
2.9 Collisional effects calculated for Maxwellian (M) versus Druyvestein (D) electron velocity distributions ........................................................................... 74
2.10 Altitude profiles of parameters employed in atmospheric chemistry model .................. 87
2.11 Equilibrium quantities from atmospheric chemistry model ................................................. 93

3.1 Test run of cylindrical particle blob propagating in uniform velocity field using fluxcorrected transport (FCT) algorithm ......................................................................................... 122
3.2 Test run of FISHPAK cylindrical Poisson solver ................................................................. 128

4.1 Vertical electric field versus position at t = 3 ms after onset of positive CG lightning discharge ......................................................................................... 133
4.2 Horizontal electric field versus position for conditions of Figure 4.1 ............................. 134
4.3 Electron temperature versus position for conditions of Figure 4.1 ..................................... 136
4.4 Altitude profiles along the simulation axis of additional quantities obtained under the conditions of Figure 4.1 ................................................................................. 138
4.5 Change in electron density from quiescent values for conditions of Figure 4.1 .................................. 140
4.6 Time history of thermal quantities associated with positive CG discharge that lowers 200 C to ground................................................................. 143
4.7 Altitude dependence of lightning-induced electric field pulses ....................................................... 144
4.8 Influence of electron density profile on simulation results .......................................................... 152
4.9 Influence of cloud charge lowered to ground by lightning stroke .............................................. 155
4.10 Influence of lightning stroke duration on simulation results ..................................................... 158
4.11 Electron temperatures throughout simulation volume at 3 ms after onset of lightning discharge ......................................................................................................................................................... 161
4.12 Effects of positive versus negative cloud-to-ground (CG) lightning ........................................... 162
4.13 Persistent effects after positive CG lightning discharge............................................................ 166
4.14 Simulation results showing downward movement of temperature depression during 40 kA, positive cloud-to-ground discharge .............................................................. 169
4.15 Simulation results showing downward-propagating discontinuity in electron density during system relaxation following a 160 kA peak current positive cloud-to-ground discharge 172

5.1 Elements of the one-dimensional attachment wave model ......................................................... 179
5.2 3-D representation of \((\bar{n}_e, \bar{n}_i, \bar{E})\) phase plane .................................................................. 188
5.3 Calculated one-dimensional solution profiles for downward propagating attachment wave in presence of - 1 V/m low altitude electric field .......................................................... 191

B.1 Influence of recombination coefficient and straw man total cluster ion density on value of \(\delta_H\), effective dehydration fraction .................................................................................. 209

C.1 Comparison of electron temperatures obtained for identical lightning discharges under the monopolar and the dipolar cloud charge models ........................................................................ 212
C.2 Comparison of simulation results obtained using the monopolar and dipolar cloud charge models for discharge parameters of Figure C.1 .............................................................................. 214
List of Tables

1.1 Comparison among lightning-induced, high-altitude, luminous phenomena ........................................... 23
1.2 Representative lightning parameters ........................................................................................................ 27

2.1 Momentum transfer collision frequencies of electrons and ions with neutrals, and cyclotron frequencies of electrons and ions at representative altitudes ................................................................. 38
2.2 Classes of charged particles involved in ion chemistry ............................................................................. 80
2.3 Charged particle production processes ..................................................................................................... 81
2.4 Charged particle loss processes ................................................................................................................ 84
2.5 Magnitudes of processes affecting the density of each class of particle .................................................... 88
2.6 Definitions of dimensionless parameters employed in fluid formulation .............................................. 101
2.7 Values of dimensionless parameters used in the fluid model .................................................................. 103

3.1 Representative duration and number of time steps for the three simulation periods ..................... 108
3.2 Definitions of simulation variables .......................................................................................................... 109

4.1 Rates of dissociative attachment and impact ionization for several temperatures at an altitude of 60 km ........................................................................................................................................... 141
4.2 Characteristics of high altitude electric field pulses at several altitudes .................................................. 145
4.3 Electron drift speed versus thermal speed for conditions obtained during the simulation .................. 148
4.4 Variation of vertical electric field pulse amplitude, maximum electron temperature $T_e$, and maximum electron density depletion $\Delta n_e$ with electron density profile ........................................................................... 153
4.5 Variation of simulation results with total cloud charge lowered to ground ........................................... 156
4.6 Variation of simulation results with lightning stroke duration ............................................................... 159

5.1 Asymptotic states of the one-dimensional system .................................................................................... 181

A.1 Scattering processes and associated energies incorporated in simulation model .............................. 202
# List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition (SI Units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>altitude of positive cloud charge center (m)</td>
</tr>
<tr>
<td>$a_n$</td>
<td>altitude of negative cloud charge center (m)</td>
</tr>
<tr>
<td>$b$</td>
<td>radial scale size of cloud charge center (m)</td>
</tr>
<tr>
<td>$c$</td>
<td>speed of light (m s$^{-1}$)</td>
</tr>
<tr>
<td>$e$</td>
<td>electronic charge (C)</td>
</tr>
<tr>
<td>$f(\vec{v})$</td>
<td>electron velocity distribution (m$^3$ s$^{-3}$)</td>
</tr>
<tr>
<td>$f_M(\vec{v})$</td>
<td>Maxwellian electron velocity distribution (m$^3$ s$^{-3}$)</td>
</tr>
<tr>
<td>$f_D(\vec{v})$</td>
<td>Druyvestein electron velocity distribution (m$^3$ s$^{-3}$)</td>
</tr>
<tr>
<td>$f(\vec{r}, \vec{v}, t)$</td>
<td>electron distribution function (m$^3$ s$^{-3}$)</td>
</tr>
<tr>
<td>$f_0(\vec{r}, \vec{v}, t)$</td>
<td>isotropic component of electron distribution function (m$^3$ s$^{-3}$)</td>
</tr>
<tr>
<td>$f_1(\vec{r}, \vec{v}, t)$</td>
<td>directional component of electron distribution function (m$^3$ s$^{-3}$)</td>
</tr>
<tr>
<td>$f_c$</td>
<td>cyclotron frequency (Hz)</td>
</tr>
<tr>
<td>$g_i(\vec{r}, T_e)$</td>
<td>ionization rate coefficient (s$^{-1}$)</td>
</tr>
<tr>
<td>$j$</td>
<td>imaginary unit</td>
</tr>
<tr>
<td>$k$</td>
<td>wavenumber (rad m$^{-1}$)</td>
</tr>
<tr>
<td>$m_\alpha$</td>
<td>mass of single particle of species $\alpha$ (kg)</td>
</tr>
<tr>
<td>$n_{e,\alpha}$</td>
<td>cluster ion density (m$^3$)</td>
</tr>
<tr>
<td>$n_e$</td>
<td>electron density (m$^3$)</td>
</tr>
<tr>
<td>$\overline{n}_e$</td>
<td>system average electron density (m$^3$)</td>
</tr>
<tr>
<td>$n_m$</td>
<td>negative ion density (m$^3$)</td>
</tr>
<tr>
<td>$n_p$</td>
<td>small positive ion density (m$^3$)</td>
</tr>
<tr>
<td>$n_e$</td>
<td>electron density (m$^3$)</td>
</tr>
</tbody>
</table>
$q_\alpha$ charge on species $\alpha$ (C)
$q_i$ collisional cross section for process $i$ ($m^2$)
$q_{D}(\vec{\sigma})$ collisional cross section for dissociative attachment of electrons to oxygen ($m^2$)
$t$ time (s)
$t_c$ duration of cloud electrification (s)
$t_l$ duration of lightning discharge (s)
$\vec{u}_\alpha$ fluid velocity of species $\alpha$ ($m \, s^{-1}$)
$\vec{v}$ particle velocity ($m \, s^{-1}$)
$z$ altitude (m)
$A(r)$ radial factor of high altitude electrostatic potential solution
$\vec{B}$ magnetic flux density (T)
$B_0$ energy of rotational transition (J)
$B(z)$ vertical factor of high altitude electrostatic potential solution
$B_C$ hydration rate ($m^6 \, s^{-1}$)
$CG$ cloud-to-ground
$\vec{E}$ electric field ($V \, m^{-1}$)
$E_{100}$ ground-level vertical electric field 100 km from discharge ($V \, m^{-1}$)
$E_f$ fair weather vertical electric field ($V \, m^{-1}$)
$G$ net rate of impact ionization ($m^3 \, s^{-1}$)
$G_{M,D}(\vec{r}, T_e)$ ionization rate computed using Maxwellian (M) and Druyvestein (D) distributions ($m^3 \, s^{-1}$)
$G_A(\vec{r}, T_e)$ dissociative attachment rate computed using Druyvestein distribution ($m^3 \, s^{-1}$)
$\vec{H}_\alpha$ heat flux vector for species $\alpha$ ($J \, m^{-2} \, s^{-1}$)
$I$ ion pair production rate due to cosmic rays ($m^3 \, s^{-1}$)
$J_0$ Bessel function of zero order
$J_f$ fair weather current density ($A \, m^{-2}$)
$J_R$ radial current density in earth-centered spherical system ($A \, m^{-2}$)
$\vec{J}$ current density ($A \, m^{-2}$)
$\vec{J}_S$ cloud source current density ($A \, m^{-2}$)
\( N \)  
total neutral density (m\(^3\))

\( N_X \)  
density of neutral species X (m\(^3\))

\( P \)  
isotropic pressure (Pa)

\( P \)  
pressure tensor (Pa)

\( Q_Z \)  
total charge in positive and negative cloud charge centers (C)

\( R \)  
reflection coefficient

\( R_E \)  
earth radius (m)

\( r \)  
fractional power lost to inelastic processes (s\(^{-1}\))

\( R_i(\tilde{r}) \)  
rate of collisional process \( i \) (s\(^{-1}\))

\( R_{i,M}(\tilde{r}) \)  
rate of collisional process \( i \) computed using Maxwellian electrons (m\(^3\) s\(^{-1}\))

\( R_{i,D}(\tilde{r}) \)  
rate of collisional process \( i \) computed using Druyvestein electrons (m\(^3\) s\(^{-1}\))

\( S_\alpha \)  
all-source particle production rate (m\(^3\) s\(^{-1}\))

\( T_0 \)  
standard temperature (K)

\( T_e \)  
electron temperature (K)

\( T_{ef} \)  
effective temperature (K)

\( V_0 \)  
earth-ionosphere potential (V)

\( \alpha \)  
ion recombination rate (m\(^3\) s\(^{-1}\))

\( \alpha_3 \)  
three-body attachment rate (m\(^6\) s\(^{-1}\))

\( \alpha_D \)  
dissociative attachment rate (m\(^3\) s\(^{-1}\))

\( \alpha_E \)  
electron-small positive ion recombination rate (m\(^3\) s\(^{-1}\))

\( \alpha_{EC} \)  
electron-cluster ion recombination rate (m\(^3\) s\(^{-1}\))

\( \beta \)  
aerosol attachment rate (m\(^3\) s\(^{-1}\))

\( \lambda \)  
detachment rate (m\(^3\) s\(^{-1}\))

\( \lambda_C \)  
dehydration rate (m\(^3\) s\(^{-1}\))

\( \lambda_\alpha \)  
eigenvalues

\( \delta_{ei} \)  
fraction of electron energy given up in momentum transfer collision

\( \delta_H \)  
effective dehydration fraction

\( \tilde{e} \)  
electron kinetic energy normalized to thermal energy
\( \varepsilon \) permittivity (F m\(^{-1}\))

\( \varepsilon_0 \) permittivity of free space (F m\(^{-1}\))

\( \Delta \varepsilon_i \) electron energy lost to collisional process \( i \) (J)

\( \mu_0 \) isotropic mobility (m\(^2\) V\(^{-1}\) s\(^{-1}\))

\( \mu_R \) reduced mobility (m\(^2\) V\(^{-1}\) s\(^{-1}\))

\( \mu_T \) transverse mobility (m\(^2\) V\(^{-1}\) s\(^{-1}\))

\( \mu_a \) isotropic mobility of species \( \alpha \) (m\(^2\) V\(^{-1}\) s\(^{-1}\))

\( \mu_\perp \) perpendicular mobility (m\(^2\) V\(^{-1}\) s\(^{-1}\))

\( \mu_{||} \) parallel mobility (m\(^2\) V\(^{-1}\) s\(^{-1}\))

\( V_m, V_{mom} \) momentum transfer collision frequency (Hz)

\( \rho \) charge density (C m\(^{-3}\))

\( \rho_f \) fair weather charge density (C m\(^{-3}\))

\( \rho_s \) cloud source charge density (C m\(^{-3}\))

\( \sigma \) total conductivity (Ω\(^{-1}\) m\(^{-1}\))

\( \sigma_+ \) positive polar conductivity (Ω\(^{-1}\) m\(^{-1}\))

\( \tau_\alpha \) mean collision time for species \( \alpha \) (s)

\( \varphi \) electrostatic potential (V)

\( \omega \) angular frequency (rad s\(^{-1}\))

\( \bar{\Gamma} \) normalized conductivity gradient (m\(^{-1}\))

\( \Lambda \) vertical component of high altitude electric field solution

\( \Phi \) electrostatic potential from numerical solution of Poisson Equation (V)

\( \Phi_{M,D}(F, T_e) \) electron energy loss rate computed using Maxwellian (M) or Druyvestein (D) distribution (W)
Chapter 1

Introduction

The electrical discharges that accompany violent storms have, since the dawn of mankind, been identified with their sensible manifestations: thunder and the flash known as lightning. Not until the early twentieth century did their electromagnetic manifestations become apparent, as the crackle of sferics on the radio. Only in the fifties and sixties did we become aware that lightning has invisible effects far beyond the atmosphere, in the form of electromagnetic whistler waves that propagate out to distances of several earth radii and back [Storey, 1953], and change the velocity of geomagnetically trapped electrons along the way [Kennel and Petschek, 1966]. Now in the nineties, with the improved sense of sight afforded by low light level video cameras, we discover that lightning gives rise to spectacular pillars of multicolored luminosity that tower above thunderstorms and reach into the lower ionosphere [Franz et al., 1990]. These features have been given the name sprites.

Groundbreaking observation of this amazing phenomenon is still underway, but it is already clear that the optical display is only one of several related, lightning-induced effects on the lower ionosphere. Among the other effects are the excitation and ionization of ambient neutrals in a huge volume above the lightning discharge [Green et al., 1996]. It is the purpose of the present work to help explain these effects of lightning by examining the flow and heating of the charged ionospheric species in response to a lightning discharge.

Briefly stated, the goal of this effort is to predict lightning-induced changes in the density, fluid velocity, and temperature of the charged constituents in the altitude range 20 - 90 km. It is intended that these predictions provide the basis for modeling the optical, infrared, and radio frequency signatures of the high altitude phenomena associated with lightning.

The importance of being able to model the densities, flows, and temperatures of ions as well as electrons stems from the great difficulty of measuring these parameters at the locations of interest. In-situ measurement techniques, using balloons and sounding rockets, are in principle capable of a complete characterization of the medium. However, balloons are limited in altitude to
a maximum of 40 km [Rosen and Hofmann, 1981a], and sounding rockets, although they traverse
the required altitudes, travel at such high velocities that they are ill-suited to study phenomena that
vary rapidly in both space and time [Ernstmeyer et al., 1993]. Of the possible remote sensing
techniques, ionosondes and incoherent backscatter radar, ionosondes provide information only
about the electron component [Bibl and Reinisch, 1978]. Incoherent backscatter radar, while
sensitive to the densities, temperatures, and drifts of both electrons and ions, is available only at a
handful of large, fixed-site facilities in the world [Evans, 1969].

Due to these limitations, we cannot expect empirical means to adequately characterize the
ionospheric response to so sporadic, unpredictable, and highly localized an event as lightning. As a
result, we must rely on detailed and accurate models of both the ambient environment and the
relevant physical processes to fully explain the phenomenon of sprites.

Although the ionospheric events we are examining take place far from our everyday world,
their importance is far from purely academic. For example, we will see that the thermal energy
deposited in the lower ionosphere after an average positive cloud-to-ground lightning discharge
approaches $10^3$ Joules. At a global flash rate of 100 per s [Uman, 1987, p 47], in 12 hours we
reach the energy of a ton of TNT (4.2E+09 J) exploding in the lower ionosphere [Mechtly, 1973].
Of course, this is in harmony with the present balance of nature, but if it is true that we are
experiencing climate changes wrought by mankind, the associated change in thunderstorm
frequency may lead to ionospheric effects that we do not yet appreciate.

Sprites have a more immediate impact as well, arising from the fact that their optical and
electromagnetic signatures, as observed by earth orbiting satellites, can mimic those of low-yield
nuclear detonations [Scott, 1995]. The unambiguous detection of nuclear detonations worldwide is
a critical Department of Defense mission. Since the end of the Cold War, the emphasis in this
mission area has shifted from supporting tactical and strategic decision making to monitoring
nuclear proliferation. A scenario which arouses extreme concern is the covert testing of small,
low-tech nuclear weapons by states intent on achieving nuclear capability. This evolving threat has
made it increasingly likely that lightning discharges and associated phenomena could either be
confused with, or used to mask, a nuclear detonation.

The present nuclear detection systems are satellite-borne and rely principally on the
impulsive optical signature (visible and infrared) of a nuclear detonation to signal an event.
Lightning discharges—in particular, a class of extraordinary lightning discharges believed to be the
cause of sprites—emit over a wide spectrum from extremely low frequencies (300 Hz to 3 kHz,
ELF) through optical and infrared, all the way to X- and gamma rays. These unusual lightning discharges have optical and IR signatures that are likely to be mistaken for those of nuclear detonations. Moreover, future sensors used for nuclear detection are expected to have greater sensitivity at optical and IR frequencies, and will include ELF/VLF and acoustic wave detectors—none of which will exclude the entire class of unusual lightning discharges. In fact, it is anticipated that lightning will be even more troublesome for the next generation of sensors, particularly as the nuclear events of concern will exhibit increasingly unpredictable signatures.

Lightning discharges thus produce phenomena in the lower ionosphere that are not only of great scientific interest, but also of potentially immense practical impact. Simulation and modeling is at present the only means of addressing key elements of the underlying physics. We are thereby motivated to develop and apply a comprehensive, self-consistent model of lightning effects on the lower ionosphere. We will introduce the effort by first presenting an overview of the observed characteristics of the newly discovered sprite phenomena and of the associated lightning. We will then discuss previous approaches to modeling these phenomena and finally distinguish the present approach.

1.1 Observations

The model which we are setting out to develop is grounded on a set of observed facts, which will not only provide a foundation for the model, but also a means to validate its results. The relevant data pertain to the sprites themselves, the medium in which they occur, and the characteristics of the causative lightning.

1.1.1 Sprites and Related Phenomena

Optical Measurements

Luminous discharges above thunderstorms were reported as long ago as 1886, when officers of the Royal Mail Ship Moselle, departing Jamaica, observed darts of light shooting vertically upward from the top of a cloud bank over the island [Toynbee and Mackenzie, 1886]. The darts achieved varying maximum altitudes, and were colored red and green, or orange and blue. However, not until the first low light level video images of the phenomenon were obtained in July 1989 [Franz et. al., 1990] did a flurry of scientific activity result—directed at elucidating the properties of the discharges and explaining their origin.
The event reported by Franz et al. [1990] was observed at a site near Minneapolis, over which there was clear air extending all the way to the location of the event, some 250 km to the north. The event itself consisted of a twin flash of light, originating at the tops of distant thunderclouds and extending vertically upward for a distance estimated to be 20 km. The flash persisted for two 17 ms television frames. The circumstance of the bulk of the thundercloud lying below the horizon at the observation point is important in that it prevents possible simultaneous lightning flashes from overwhelming the high altitude luminosity. Since no time-correlated sferic observations were available, it is not known whether the event was simultaneous with a lightning discharge.

Subsequent observations of the so-called “stratospheric lightning” were made from the Space Shuttle Discovery in October 1990 [Boeck et al., 1992]. Like the Minneapolis observations, these were essentially fortuitous. Data from a Space Shuttle payload bay video camera, which was used routinely to measure cloud top optical emissions produced by lightning, captured nine images of upward vertical discharges. These discharges appeared to begin at the altitude of the cloud tops and to extend upwards for approximately 35 km.

By 1993 these reports had stirred enough interest that the first systematic campaigns to detect the high altitude luminous effects of lightning were mounted. Measurements made from the NASA DC-8 Airborne Laboratory in July 1993 [Sentman and Wescott, 1993] were the first to record a large number (at least 19) of upper atmospheric optical flashes in association with a single thunderstorm. An all-sky camera atop the fuselage recorded flashes as the aircraft flew at an altitude of 12 km, skirting a severe storm system in Iowa, Kansas and Nebraska. The resulting observations established that the events were of short duration (≤ 17 ms), were as bright as strong aurora (10 - 50 kR), had diameters in the range 10 - 50 km, and extended to altitudes averaging 60 km, with some as high as 100 km. Although the upper atmospheric flashes frequently coincided in time with conventional lightning, it was not possible to correlate individual events with specific lightning discharges as recorded by the National Lightning Detection Network (NLDN).

Nevertheless, the authors point out that based upon the relative frequency of negative versus positive cloud-to-ground (CG) lightning discharges detected by the NLDN during the observation period (roughly 9:1), there was one high altitude flash for every 200 to 400 negative CG discharges or one for every 20 to 40 positive CG discharges.

A campaign of nightly, ground-based monitoring of the stratosphere above distant thunderstorms in July and August of 1993 [Lyons, 1994] recorded more than 600 instances of
cloud-to-stratosphere lightning using low-light video. The camera was stationed on Yucca Ridge, 20 km northeast of Fort Collins, CO, with a clear view of the plains to the east. The images revealed massive columns of luminosity with bright tops and hanging tendrils. Accompanying naked-eye observations revealed one such object to have the red-orange color of salmon. The author suggested that the events were associated with huge, mature convective storm systems measuring at least 100-200 km on a side that had been active for several hours prior to the onset of high altitude flashes. The flashes were clustered in regions where positive cloud-to-ground lightning strokes were relatively common. At this time the descriptive name “red sprite” was coined for the phenomenon, at the suggestion of Dr. David Sentman of the University of Alaska.

Additional low-light video observations taken in July and September of 1994 cemented the correlation between sprites and positive CG lightning discharges [Boccippio et. al., 1995]. Based upon a detailed comparison of the times of occurrence of the observed sprites, NLDN-derived CG stroke times, and the onset times of certain extremely low frequency (ELF) electromagnetic transients detected in Rhode Island, a one-to-one correspondence was established between almost every sprite and a companion cloud-to-ground lightning discharge. It was thereby determined that of 42 sprites detected on 12 July 1994, 36 were associated with positive CG strokes, and 45 out of 55 sprites on 7 September 1994 were likewise associated.

Since 1994, additional campaigns of coordinated ground- and aircraft-based measurements over the American Midwest have confirmed and refined initial observations of the optical properties of red sprites. The established optical properties of sprites may be summarized as follows:

1. Sprites attain an average maximum altitude of 88 km. They are predominately red, occur in clusters of two or more, and possess a distinctive visible morphology consisting of (in order of decreasing altitude) a faint, wispy hair-like structure, between 74 and 88 km; a region of maximum brightness, between 66 km and 74 km, known as the head; a dark band, or collar; and tendrils extending from 66 km down to 40 - 50 km [Sentman et. al., 1995].

2. Sprites last for approximately 3 ms [Sentman and Wescott, 1995] and are delayed from the associated lightning flash and VLF sferic by 3 -30 ms [Fukunishi et al., 1996].

3. Sprites are strongly correlated with positive cloud-to-ground (CG) lightning strokes exhibiting anomalously large peak return stroke currents in the range 100 kA - 200 kA [Boccippio et. al., 1995].
The continuing efforts to observe sprites have uncovered several additional phenomena: blue jets, blue starters, and elves. Blue jets were first observed over southern Arkansas on the night of 1 July 1994 [Wescott et al., 1995]. They are distinguished from sprites by their predominately blue color, lower altitudes, and longer duration. The jets persist for as long as 350 ms, or up to 20 video frames, so that their evolution is apparent from the video footage. This is a considerable contrast to the situation for sprites, which vanish in less than a frame period—making their evolution inscrutable. The jets are seen to emanate from the cloud tops and grow vertically upward into a narrow, 15° cone at an average speed of 98 km/s. Terminal altitude is typically 40 km. The blue jets are associated not with CG lightning of either polarity, but instead with what was then considered to be an unusual form of lightning that extends upward a short distance from the cloud top.

Subsequent analysis revealed that the apparent upward lightning discharges were in fact representative of yet another distinct new phenomenon that was given the name “blue starter” [Wescott et al., 1996]. The starters are much shorter in vertical extent than the jets, ranging from a mean bottom altitude of 17.7 km (roughly even with the cloud tops) to a mean top altitude of 20.8 km. Clear examples have been found in which blue starters are seen to rise directly from the anvil of a thundercloud. The starters were never observed to be simultaneous with either variety of cloud-to-ground lightning, generally occurring in quiet periods between lightning discharges. They were, however, concentrated in small geographic regions (~20 km²) experiencing high rates of negative CG discharges as well as large hail (5 cm diameter). Rates of negative CG lightning dropped markedly subsequent to the occurrence of a blue starter.

Finally, as a result of the SPRITES ’95 observation campaign, large, extremely short-lived diffuse flashes were observed to precede the occurrence of sprites [Fukunishi et al., 1996]. These events, now called “elves,” appeared at the same and somewhat higher altitudes than sprites (75-105 km), and exhibited a horizontal extent of 100-300 km. The elves appeared within 500 μs of the onset of the associated lightning discharge, and persisted for less than 1 ms. They had thus disappeared by the time of the subsequent sprites, which appeared 3 - 30 ms later. It was also determined that the luminous layer of elves moved downward at a speed of 0.1c to 0.3c.

Characteristics of the new lightning-related phenomena discovered in the nineties are summarized in Table 1.1. Figure 1.1 displays the typical altitudes of the various phenomena in relationship to the atmospheric and ionospheric layers. Sprites reside principally in the mesosphere, and in the lowest reaches of the ionosphere.
Figure 1.1: Altitude ranges of lightning-excited phenomena in relation to ionospheric and atmospheric layers.
<table>
<thead>
<tr>
<th></th>
<th>RedSprites</th>
<th>Blue Jets</th>
<th>Blue Starters</th>
<th>Elves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude extent</td>
<td>40 - 90 km</td>
<td>20 - 40 km</td>
<td>18 - 21 km</td>
<td>75 - 105 km</td>
</tr>
<tr>
<td>Horizontal extent</td>
<td>20 - 50 km dia</td>
<td>3 km dia base 10 km top</td>
<td>&lt; 1 km</td>
<td>100- 400 km dia</td>
</tr>
<tr>
<td>Shape</td>
<td>three bright patches separated by dark bands, vertical striations</td>
<td>upward flaring cone 15° full width</td>
<td>vertical, spark-like</td>
<td>diffuse blob</td>
</tr>
<tr>
<td>Time of onset</td>
<td>3 - 30 ms after lightning onset</td>
<td>quiet periods</td>
<td>quiet periods</td>
<td>&lt; 500 µs after lightning onset</td>
</tr>
<tr>
<td>Duration</td>
<td>3 ms</td>
<td>200 - 300 ms</td>
<td>NA</td>
<td>&lt; 1 ms</td>
</tr>
<tr>
<td>Relationship to lightning</td>
<td>follow positive CG lightning</td>
<td>during periods of intense negative CG lightning</td>
<td>during periods of intense negative CG lightning</td>
<td>follow positive CG lightning</td>
</tr>
</tbody>
</table>

Table 1.1: Comparison among lightning-induced, high-altitude, luminous phenomena

As an extension of the optical measurements of sprites, the observed sprite spectra have been analyzed to estimate the temperature of electrons within the sprites. Green et al. [1995] used optical spectra measured in July 1995 to estimate the relative populations of the vibrational states of molecular nitrogen in several sprites. The observed spectra were corrected for the effects of atmospheric transmission, and the results were compared with synthetic spectra that depended upon the vibrational state populations. The squared difference between the two sets of spectra was minimized by adjusting the populations of the vibrational states. From the resulting relative populations, the authors inferred the temperature of an assumed Maxwellian distribution of electrons. From this analysis, it was found that the observed optical spectra were consistent with electron temperatures between 0.4 and 2.0 eV. Since quiescent electron temperatures at these altitudes are on the order of 0.02 eV [Champion et al., 1985], this suggests that significant electron heating is related to the sprites. In fact, according to the analysis of Green et al. [1996], electron heating is responsible for sprite luminosity.

Electric Field Measurements

Since electric field measurements must be made in situ, they are considerably harder to come by than optical measurements. Nevertheless, electric fields have been observed in the mesosphere above thunderstorms, at various altitudes and using various platforms. In 1986, a NASA U-2 research aircraft was used to measure the vertical electric field above summertime air mass thunderstorms as part of a large meteorological experiment campaign undertaken in the
Tennessee Valley Region of the United States [Blakeslee, et al., 1989]. The quasi-electrostatic, vertical electric field at 20 km altitude was seen to reach peak values of 7 kV/m. Abrupt discontinuities in the electric field, coincident with lightning discharges, had a typical magnitude of 4 kV/m and tended to reduce the absolute magnitude of the field. No high speed transient fields were recorded.

An experiment carried out in July 1988 detected electric fields from two sounding rockets (Thunder Hi and Thunder Lo) that overflowed a thunderstorm in the Atlantic Ocean off the coast of Wallops Island, Virginia [Kelley et al., 1985, 1990]. The instrument flown was able to digitize three-axis electric field signals at rates up to 20 kilosamples per second. More than a hundred storm-related electric field pulses were detected during the flight. A representative example was measured at an altitude of 142 km by Thunder Hi and at an altitude of 88 km by Thunder Lo. It was directed downward, parallel to the earth’s magnetic field (inclined approximately 70° from horizontal). The pulse had an amplitude of approximately 30 mV/m at 142 km altitude, and approximately 10 mV/m at 88 km. The electric field transient lasted for at least 10 ms at 142 km, and for at least 15 ms at 88 km. It followed within two milliseconds an optical lightning flash detected by the payload. Given the short time delay from the optical flash, the disturbance associated with the electric field pulse must have traveled at a considerable fraction of the speed of light.

Pulses of comparable magnitude were observed by an electric field experiment on board the DE 2 satellite as the spacecraft overflowed Hurricane Debbie at an altitude of 295 km in 1982 [Burke et al., 1992]. The hurricane was located in the Atlantic Ocean. In this case, the peak electric field was approximately 40 mV/m, and directed perpendicular to the earth’s magnetic field. Due to limitations of the instrument, the duration of the pulse can be characterized only as shorter than one second. The electric field transient was accompanied by a burst of field-aligned electrons with energies of roughly 1 keV. The burst lasted 30 ms. While the two phenomena appear to be related, the direction of the electric field pulse suggests that it did not directly accelerate the electrons.

It must be noted that while these in situ transient electric field measurements are known to be associated with lightning discharges, it is not known whether those lightning discharges coincided with the occurrence of a sprite or jet. As a result, it is not certain that any in situ observations made so far actually apply to the environment just before and during a sprite event. Moreover, it has not been possible to directly measure the densities, velocities, and temperatures of
the charged constituents at the time and location of a sprite, and remote measurements have been limited to electron temperature. Even if balloons or sounding rockets were instrumented to measure these properties, the fleeting and unpredictable nature of the lightning-induced phenomena would make it difficult to obtain a large number of observations. We thus arrive at the need for modeling and simulation efforts to complete the picture of lightning-induced phenomena, and to predict some observable attributes for purposes of model validation.

1.1.2 Lightning Phenomenology

We have pointed out that red sprites are exclusively associated with positive cloud-to-ground lightning, while blue jets are associated with lightning-producing storms in a more general sense. In order to provide the necessary context for the sprite observations, we must thus examine lightning itself.

Lightning culminates the long and as yet poorly understood process of thundercloud electrification. The numerous models proposed to account for the separation of charge within the cloud are not of interest here, only the observed facts regarding the time evolution of the cloud charge. Electrification commences with the appearance of precipitation within the cloud and typically progresses for several minutes until the first lightning flash [Israel, 1973, p. 523]. The rate of electrification and the time to first flash depend upon details of the precipitation development in particular storms and are therefore somewhat variable. After the initial lightning flash, electrification resumes and an individual storm cell exhibits subsequent flashes at a typical rate of one per minute [Smith, 1958].

The separation of charge within the thundercloud results, roughly speaking, in the creation of a vertically aligned electric dipole. In a typical summertime thundercloud, negative charge accumulates in a region centered at an altitude of 7 km, while positive charge accumulates about a center that ranges in altitude from 10 to 15 km [Krehbiel, 1986, p. 96].

Lightning can be contained entirely within the cloud (intra-cloud, or IC, lightning), or can extend from the cloud to the earth’s surface (cloud-to-ground, or CG, lightning). IC lightning generally outnumbers CG lightning by a factor of five or ten [Krehbiel, 1986, p. 105]. CG lightning is further divided between positive and negative varieties, depending upon the sign of the charge lowered to the earth as a result of the discharge. Of the two, positive CG lightning is relatively infrequent during the summer, accounting for only 4% of all lightning flashes in one study [Orville et al., 1983].
A negative CG lightning flash begins in the cloud with a preliminary electrical breakdown that initiates the so-called stepped leader, which is a highly branched discharge that propagates horizontally and downward in spurts of 30 to 90 m every 20 to 100 μs [Krider, 1986, p. 31]. The stepped leader typically reaches within striking distance of the ground 100 ms after the preliminary discharge. Although the stepped leader delivers some negative charge to the ground, the main current-carrying event is the return stroke that immediately follows it. The return stroke is a wave of ionization that propagates upward along the conductive channel left by the leader at speeds of roughly 1.0E+08 m/s, neutralizing both the lightning channel and a significant amount of cloud charge. Peak current for a first negative return stroke has a median value of 30 kA, and a 95th percentile value of 80 kA [Uman, 1987, p. 124]. By the end of the first return stroke, the net effect has been that a median value of 5 C (95th percentile value of 25 C) of negative charge has been lowered to ground. The first return stroke is followed after a pause of 40 to 80 ms by a subsequent leader and return stroke along the original lightning channel. Most negative CG flashes involve 2 to 4 return strokes. The duration of each stroke is on the order of 100 μs.

Positive CG flashes are associated with large vertical shear in the horizontal wind, which displaces the upper positive charge center of the thundercloud horizontally from the lower negative charge center [Rust, 1986, p. 42]. The positive CG lightning then connects the upper positive charge center and the earth. The evolution of a positive flash generally follows the same outline as a negative one, with a downward propagating leader followed by an upward return stroke [Fuquay, 1982]. In addition to the difference in polarity, positive CG lightning is distinguished by much higher currents, and much greater total charge transferred. Median peak current in positive return strokes is 35 kA (95th percentile value of 250 kA), and median total charge lowered to earth is 80 C (95th percentile value of 350 C) [Uman, 1987, p. 124]. Positive CG flashes typically involve only a single return stroke, and typically manifest a period of continuing current after the stroke. Rise times (10% to 90%) for the return stroke current are less than 10 μs, while decay times are of order 1 ms.

Table 1.2 shows representative parameter values for both positive and negative CG lightning. These values, derived from the above discussion, will form the basis of the lightning models employed in this effort.
<table>
<thead>
<tr>
<th></th>
<th>Average Neg CG</th>
<th>Extreme Neg CG</th>
<th>Average Pos CG</th>
<th>Extreme Pos CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>peak current (kA)</td>
<td>30</td>
<td>80</td>
<td>35</td>
<td>250</td>
</tr>
<tr>
<td>charge lowered (C)</td>
<td>5</td>
<td>25</td>
<td>80</td>
<td>350</td>
</tr>
<tr>
<td>rise time (μs)</td>
<td>5</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>decay time (μs)</td>
<td>500</td>
<td></td>
<td>1000</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.2: Representative lightning parameters.

We have presented a framework of observed facts regarding sprites, jets, and their causative lightning. This framework will guide our construction of a simulation system and provide a benchmark for the results.

1.2 Previous Simulation Approaches

In the several years that followed the initial Minnesota sprite observations [Franz et al., 1990], considerable effort has been spent in modeling lightning-induced phenomena in the mesosphere/lower ionosphere. The models that have emerged so far have taken three principal approaches: (1) a purely local, kinetic theory based upon runaway air breakdown, (2) quasi-local theories that compute the response of the local medium to an *ad hoc* electric field pulse, and (3) a global, quasi-electrostatic theory accounting for the effects of cloud height charge removal on the medium at all altitudes up to the base of the ionospheric E layer.

Gurevich et al. [1992] and Roussel-Dupre et al. [1994] elaborated on a suggestion by Wilson [1924] that runaway breakdown of air could occur in a region of intense electric field above a thunderstorm. This model assumes that cosmic ray secondaries form a seed population of electrons upon which the electric field acts to create a population of primary avalanching electrons. The primary electrons generate bremsstrahlung (x- and gamma rays), and create a population of secondaries which thermalize quickly, and thereby form a drifting, high temperature electron gas. The secondary population is directly responsible for the sprites. In this kinetic treatment, the electron populations evolve according to the Boltzmann equation.

Taranenko et al. [1993] and Rowland et al. [1995] have suggested that strong electromagnetic pulses from lightning could interact with the lower ionosphere to produce sprites. Taranenko et al. [1993] solve the Boltzmann equation for the electron distribution function locally at each altitude, including all manner of inelastic collisional losses. Rowland et al. [1995], on the other hand, employ empirical swarm data to calculate electron fluid properties directly from the
electric field amplitude. Both groups then introduce an *ad hoc*, 100 \( \mu \)s long, electromagnetic pulse at an altitude of 70 km or higher to simulate the effect of lightning. Considerable enhancement of electron density and temperature is predicted by both groups.

Finally, Pasko et al. [1995] present a self-consistent treatment of the effects of quasi-electrostatic fields due to positive CG lightning discharges. Their simulation volume is cylindrically symmetric and extends essentially from the earth’s surface all the way to the base of the ionosphere, taken to be at 80 km. Lightning is modeled as the sudden removal of cloud charge from the 10 km to 20 km altitude region. The resulting quasi-electrostatic disturbance heats ambient electrons at 70 km - 85 km altitudes, and alters the electron density there by a combination of dissociative attachment and impact ionization. In this model, electric fields evolve according to Poisson’s Equation with an external source term to represent cloud charge; current density obeys Ohm’s Law with conductivity calculated directly from the local electric field using empirical swarm data; and charge density evolves according to the equation of current continuity. The model is solved by a finite element calculation and predicts optical emissions roughly consistent with sprite observations.

Because the runaway air breakdown theory of Gurevich et al. [1992] and Roussel-Dupre et al. [1994] provides solutions only for a homogeneous medium under the influence of a uniform electric field, it applies to the study of the sprite environment only in the local approximation. As a result, it is not directly comparable to the approach taken here, which accounts self-consistently for the spatial variation of the medium, and we will not discuss it further. This leaves the electromagnetic pulse (EMP) theories and the quasi-electrostatic (QE) theories as comparable examples of previous modeling approaches. The two classes of theories (EMP and QE) treat two different and largely independent aspects of the problem. The EMP theories address the extremely rapidly propagating effects that take place within the first several hundred microseconds after a lightning discharge, while the QE theories examine the slower, longer-lasting effects of cloud-height charge removal. Both approaches are necessary elements in a complete picture of lightning-induced mesospheric effects.

**EMP Theories**

The first theory of the interaction of lightning-induced electromagnetic pulses with the lower ionosphere is due to Rodriguez et al. [1992]. This formulation models a horizontal intra-cloud (IC) discharge as an infinitesimal horizontal dipole 5 km above the ground. The ordinary
and extraordinary magnetoionic components of the radiated wave are propagated upward along a
range of ray paths and the power dissipated in the medium is calculated at altitude intervals of 0.1
km. The change in electron-neutral collision frequency due to the associated heating is tracked.
The number density of secondary electrons created per ionization period is estimated by dividing
the total energy in the high-energy tail of the electron distribution by the average ionization energy
of the ambient neutrals. This ionization does not self-consistently affect the reflection height of the
pulse.

Rodriguez et al. [1992] find that even relatively weak pulses produce peak electron
temperature enhancements of a factor of ten over the quiescent temperature, while intense lightning
produce a several hundred-fold increase. The customary means of characterizing stroke intensity is
the magnitude of the vertical electric field produced on the ground at a distance of 100 km from the
discharge, $E_{100}$. In the context of Rodriguez et al. [1992], a weak discharge has $E_{100} = 1$ V/m,
while an intense one has $E_{100} \geq 20$ V/m. The fractional increase in electron density due to
ionization approaches unity only for the most intense discharges ($E_{100} = 50$ V/m) studied, under
circumstances of tenuous ambient electron density, which allow pulse penetration to higher
altitudes ($\sim 100$ km) where the electron-neutral ionizing collision period is longer. For denser
ambient electron profiles, peak ionization occurs at altitudes as low as 90 km. Rodriguez et al.
[1992] also compare the effects of horizontal and vertical discharges and find that horizontal
discharges produce greater maximum heating and ionization, while vertical discharges create
disturbances of larger horizontal extent.

The next development in the modeling of electromagnetic pulses was to self-consistently
account for the lowering of the ionospheric reflection height during the passage of the pulse
[Taranenko et al., 1993]. In their approach, the medium is divided into a pile of horizontal slabs
in the altitude range 70 to 120 km. In each slab, the Boltzmann equation is solved for the local
electron velocity distribution function in the presence of a locally uniform electric field. The
electron population is taken to be at high temperature, so that the thermal velocity greatly exceeds
any drift velocity, and the electron distribution function may be split into a part, $f_0(\vec{V}, F, t)$,
which depends only on the magnitude of the velocity, and a directional part $\frac{\vec{V}}{|\vec{V}|} \cdot \vec{f}_1$. $f_0$ and
$\vec{f}_1$ evolve under the influence of a time-varying external electric field, a static magnetic field, and
elastic and inelastic collisions. The inelastic collisional processes taken into account in this

29
formulation are rotational, vibrational, and optical transitions, as well as dissociation, dissociative attachment, and ionization.

The resulting equations are solved numerically, starting from an initial Maxwellian velocity distribution. It is found that the average energy of the electrons reaches an equilibrium value within 10 µs of the onset of the electric field. Since lightning-induced electric fields vary slowly over 10 µs time scales, Taranenko et al. [1993] solve the Boltzmann equation in the steady state. The resulting quasi-equilibrium values of current density, ionization and attachment are placed in a lookup table as functions of instantaneous electric field and altitude.

The effect of lightning is introduced as a linearly polarized, 100 µs long, single-cycle electric field pulse, launched at an altitude of 70 km. Propagation of this pulse is followed through successive slabs in a one-dimensional simulation. Maxwell's equations are solved in each slab using the electric field-dependent properties of the medium obtained from the quasi-equilibrium lookup table described above. The medium not only evolves as a result of the passage of the electric field pulse, but also self-consistently affects the propagation of the pulse. The system is allowed to develop for 260 µs.

Taranenko et al. [1993] show results for an electric field pulse of peak amplitude 25 V/m at 70 km initial altitude. As the pulse propagates upward through an altitude of 83 km, electron temperature is elevated to the level at which the process of dissociative attachment dominates, and an electron density depletion of 4% is observed. At an altitude of 90 km, the maximum electron temperature of 6 eV occurs. At this temperature, the energy of a sizable fraction of the electron population exceeds the ionization threshold of the ambient neutrals, and electron density is enhanced by 27%. This combination of effects leads to a sharpening of the positive electron density gradient near 90 km altitude. Finally, the pulse is strongly attenuated above 95 km.

The effects of multiple-stroke lightning flashes are compared to those of isolated strokes. In the context of Taranenko et al. [1993], a multiple stroke translates into twenty or more electric field pulses in the short period of 2 ms. The electron density enhancements associated with the individual pulses are cumulative over this short period and produce a several-fold total increase in electron density versus ambient conditions.

In subsequent contributions [Rowland et al., 1995] to the theory of electromagnetic pulse interaction with the upper atmosphere, the Boltzmann Equation solution for electron temperature was replaced by the use of tabulated swarm data. Swarm experiments measure the drift speed and ionization rate of electrons in air as a function of applied electric field. Such measurements
inherently take into account the details of air chemistry and of the electron velocity distribution—at least for the case of relatively few electrons in a dense neutral background. In effect, a lookup table is used to obtain the electron current and net ionization rate, just as was done by Taranenko et al. [1993], but now the tabulated values are obtained empirically instead of theoretically.

The Rowland et al. [1995] simulation launches from a low altitude the electromagnetic pulse that would result from the horizontal lightning current corresponding to a given value of $E_{100}$. The current waveform is a decaying 10 kHz sinusoid, of which the high frequency component dominates in the far field. The far field vacuum result is used for the radiated fields up to an altitude of 60 km. Above this altitude, the medium is characterized by Hall, Pedersen, and parallel conductivities in the form of a magnetoconductivity tensor.

Capitalizing on the result of Taranenko et al. [1993] that electron temperature saturates at 6 eV within microseconds during a strong electric field pulse, Rowland et al. [1995] calculate the electron-neutral collision frequency at a fixed electron temperature of 6 eV. Changes in conductivity are then purely the result of changes in the plasma density produced by ionization.

The principal result of the Rowland et al. [1995] effort is that lightning discharges above a threshold intensity ($E_{100} = 20$ V/m) produce sufficient ionization in the lower D region to reflect the trailing portion of the electric field pulse. In this sense, the lower ionosphere acts as a nonlinear switch that saturates above a threshold. For comparable cases below the threshold ($E_{100} = 11$ V/m corresponds to 17 V/m electric field at 70 km altitude), the Rowland et al. [1995] peak electron density enhancement is in reasonable agreement with the results of Taranenko et al. [1993]. For stroke intensities above threshold, Rowland et al. [1995] predict huge values of peak ionization. For example, $E_{100} = 60$ V/m produces density enhancement by a factor of roughly 400. Stronger pulses produce breakdown at lower altitudes as a result of pulse reflection from the lowered ionosphere.

Subsequent work [Rowland et al., 1996] compared the effects of vertical and horizontal lightning discharges. For a given value of $E_{100}$, it is found that horizontal current produces greater maximum ionization than does vertical current. However, larger horizontal current than vertical is required to produce the same value of $E_{100}$. Since the current is the more fundamental quantity, it is therefore ambiguous whether an average horizontal lightning stroke would be more or less effective at producing atmospheric breakdown than an average vertical lightning stroke. The only clear result is that the horizontal portion of an extended lightning discharge—the so-called “spider lightning”—makes an important contribution to the ionization of the upper atmosphere.
Quasi-electrostatic Theories

Pasko et al. [1995] were the original proponents of quasi-electrostatic effects to explain sprites. They represent the cloud charge distribution in both a monopolar and a dipolar configuration. The monopolar case consists of a spherical, Gaussian distribution of positive charge at an altitude of 10 km. In the dipolar case, this is augmented by a spherical, Gaussian distribution of negative charge at 5 km altitude. In both cases, the electrification period is modeled by an exponential growth in time of the magnitude of the charge density throughout the cloud. The positive CG lightning discharge itself is represented by an exponential decrease in the magnitude of the positive charge density. The time constant used for the electrification is 0.5 s, and for the discharge, 1 ms. It is found that the monopolar and dipolar configurations produce identical effects at altitudes above 55 km.

The quasi-electrostatic formulation of the problem [Pasko et al., 1997] entails the coupled current continuity and Poisson equations:

$$\frac{\partial (\rho + \rho_s)}{\partial t} + \nabla \cdot (\overrightarrow{J} + \overrightarrow{J_s}) = 0$$  \hspace{1cm} (1.1)

$$\nabla \cdot \overrightarrow{E} = \frac{(\rho + \rho_s)}{\varepsilon_0}$$  \hspace{1cm} (1.2)

where \(\rho\) is charge density, and \(\rho_s\) and \(\overrightarrow{J_s}\) are source charge and current density, respectively.

The system is closed by computing \(\overrightarrow{J}\) in terms of \(\overrightarrow{E}\) using swarm data. Equations (1.1) and (1.2) are solved to determine the electric field subject to the boundary conditions of perfectly conductive surfaces on the cylindrical wall and the top and bottom of the cylindrical simulation volume. The simulation is undertaken in a cylindrically symmetrical geometry, in a volume that extends from the ground to an altitude of 90 km, and to a radius of 60 km. The cloud charge centers are located on the cylinder axis.

Given the calculated electric field, additional swarm data is used to estimate the coefficients of electron-neutral impact ionization, \(\nu_I\), and dissociative attachment of electrons to \(\text{O}_2\) molecules, \(\nu_A\). The local electron density then evolves in time at each point in space according to

$$\frac{dn_e}{dt} = (\nu_I - \nu_A) \cdot n_e$$  \hspace{1cm} (1.3)
The resulting electron density, coupled with the electron mobility obtained as a function of electric field from swarm data, determines the conductivity at each point in the simulation volume. The self-consistent variation of conductivity with applied electric field is a significant innovation associated with the *Pasko et al.* [1995] approach.

The fundamental result of the quasi-electrostatic model [*Pasko et al.*, 1997] is that lightning removes charge from the atmosphere-mesosphere system on time scales much shorter than the local dielectric relaxation time of the medium at altitudes up to 60 km. Since the medium responds so slowly, the lightning discharge is equivalent to the sudden placement of unshielded charge in the system. The result is a pulse of electric field at all altitudes above the cloud, with amplitude and duration that decrease with altitude, as conductivity increases. For example, a discharge involving the removal of 200 C of positive charge produces a pulse of -400 V/m amplitude and 15 ms duration at an altitude of 60 km, but only -70 V/m and 1 ms at 80 km.

In a manner similar to the electromagnetic pulse scenario [*Taranenko et al.*, 1993], the quasi-electrostatic results predict increasing electron temperatures with altitude in the 60 km to 80 km altitude range [*Pasko et al.*, 1997]. The result is a switch from dissociative attachment to ionization with increasing altitude—sharpening the electron density profile. Peak electron temperatures are in the neighborhood of 2 eV to 5 eV.

*Pasko et al.* [1997] also consider the differing effects positive versus negative cloud-to-ground discharges. It is found that +CG discharges produce several times stronger electric fields at high altitude (60 - 75 km) than do -CG discharges. The authors explain that the root cause of this effect is the fact that in their model, charge is removed from a higher altitude in the +GC case (10 km) than the -CG case (5 km). They find, for example, that the effects of removing 100 C of charge from 10 km is approximately the same as removing 50 C from 20 km. However, in the real world positive and negative CG discharges often remove charge from the same altitudes [*Boccippio et al.*, 1995].

1.3 Distinction of the Present Approach

Of all the previous modeling efforts, the work of *Pasko et al.* [1997] bears greatest similarity to the present approach. However, the present approach incorporates fundamental physical effects that were ignored in previous work. We will see that direct knowledge of electron temperature and particle fluxes, and a detailed accounting of elastic and inelastic electron-neutral
collisional processes provide otherwise unobtainable insights into the behavior of the mesosphere subsequent to a lightning discharge.

First and foremost, the semi-empirical model of Pasko et al. [1997] does not account for the flow of charged species under the influence of the quasi-electrostatic field. Electron mobility is estimated as a function of electric field through the use of swarm data, and electron density evolves in place due to impact ionization and dissociative attachment. Electron mobility and density then determine the conductivity of the medium. In the Pasko et al. [1997] formulation, neither the electrons nor the small ions that account for the low altitude conductivity are allowed to flow in response to the electric field. In our model, by contrast, the densities of three component species (electrons, and positive and negative ions) respond to macroscopic flow as well as ionization, recombination, dissociative attachment, and a great many chemical processes. This allows us to examine the large scale relocation of ions over periods of 10 s or more. As a result, we are able to follow the system over much longer periods than Pasko et al. [1997] and to predict flow-related phenomena, such as charge separation, that are beyond the scope of their formulation.

Moreover, Pasko et al. [1997] never explicitly determine the electron temperature, and instead relate electron mobility directly to electric field through the use of swarm data. The model presented here provides electron temperature by solving the electron energy transport equation, including the effects of the electric field and both elastic and inelastic collisions. The flow of electric field energy into molecular and electronic excitations is tracked accurately and in detail. This will allow us to make an estimate of bulk heating of the neutral mesosphere due to lightning discharges.

A further distinction of the present approach is the incorporation of a nonlinear, analytical model of attachment wave propagation to elucidate the properties of the non-neutral region created by a quasi-electrostatic field pulse. From the simulation of Pasko et al. [1997], it is not possible to obtain expressions for the speed at which the quasi-electrostatic fields penetrate the overlying medium, nor is it possible to investigate the fine scale of spatial variation in the quasi-electrostatic field discontinuity. These limitations are overcome in the present approach, in which a one dimensional, nonlinear fluid model is used to examine the balance among diffusion, advection, and drift due to electrostatic fields created by charge separation. This model is an extension of work by Turcotte and Ong [1968] to include the effects of temperature variation across the shock front.
Chapter 2

Three-dimensional Simulation Model

The effort described here comprises two major models. One is a comprehensive, multiple-fluid simulation that calculates the macroscopic properties of the medium in the real, physical space volume between the ground and the ionosphere, on the actual time scales of the cloud electrification and lightning discharge. The other is a three-fluid, one-dimensional, nonlinear model of attachment wave propagation that focuses on phenomena in the thin transition region between the large-magnitude, low-altitude electric field and the essentially field-free overlying medium. The one-dimensional model is used to amplify and elucidate the results of the three-dimensional simulation, and is in this sense subsidiary to it. We therefore begin with a complete discussion of the comprehensive, three-dimensional simulation, and save discussion of the attachment wave model for Chapter 5.

2.1 Overview and Geometry

The purpose of the three-dimensional simulation is to provide spatially realistic predictions that are suitable for comparison with experiment. The initial conditions are meant to mimic the natural, quiescent environment as closely as possible. The external stimulus to the system consists of the simulated lightning stroke itself. The resulting effects on the electrostatic field, on the particle densities and temperatures, and on the total conductivity are all meant to be solved self-consistently. The overall package is a system of six different component models, all operating simultaneously to account for the variety of physical effects involved.

The overarching model is the multiple-fluid model that computes changes in particle and charge density as a result of particle fluxes and collisional and chemical effects. The fluid model also computes electron temperature as a function of electric field, particle fluxes and densities. Supporting the fluid calculation is an electric field model that combines a low altitude Poisson solution with a high altitude quasi-neutral solution to improve accuracy and computability. Also supporting the fluid model is a lightning model that provides the spatial distribution and temporal
evolution of the cloud charge. A collisional effects model includes elastic electron-neutral collisions as well as inelastic processes involving the excitation, ionization, and dissociation of ambient neutral species. A chemical effects model accounts for the many competing processes of background ionization, electron attachment, hydration, dehydration, and recombination. Finally, supporting all the other calculations is a model of the ambient medium, including altitude profiles of the neutral densities, and of the electron and small ion densities. Figure 2.1a shows a schematic of the inter-relation among the several models.

Although the system is three-dimensional, we will assume cylindrical symmetry in order to take advantage of the computational efficiency afforded by having only two independent spatial coordinates. The cylindrical axis of the system is labeled $z$ and directed vertically upward. It extends from the ground up to an altitude of 90 km. The radial coordinate is $r$, extending to a radius of 64 km, and the azimuthal coordinate is ignored. The $(r, z)$ system is illustrated in Figure 2.1b.

We justify the assumption of cylindrical symmetry in three steps. First, all features of the quiescent ambient environment are essentially uniform in the horizontal directions over the 128 km diameter of the system. Thus, the ambient environment varies only with $z$. Second, the spatial distribution of charge within the cloud can realistically be modeled as symmetric about a vertical axis, which is then taken to define the $z$-axis of the system. Cloud charge is thus a function of only $r$ and $z$. Finally, the effects of the earth’s magnetic field can be ignored, so that the response of the system involves only radial and vertical flows of particles. This is realistic in the highly collisional environment under consideration because the collision frequencies of the charged particles with neutrals, $\nu_m$, greatly exceed the cyclotron frequencies, $f_c$, of the charged particles, which are therefore unable to execute organized cyclotron motion. Frequent collisions impede the $\vec{E} \times \vec{B}$ drift of the particles as well as facilitate their random walk in the direction of the electric force. Under such circumstances, the background magnetic field does not influence the particle dynamics. Representative values of the frequencies in question are displayed in Table 2.1.
Figure 2.1: Simulation overview and geometry. (a) Schematic representation of comprehensive lightning effects simulation. $S_\alpha$ represents the rate of change of the density of charged species $\alpha$ due to electrification and lightning discharge. $\tau_e, g_e, R_i$ are electron-neutral mean collision period, net impact ionization, and volume rate of inelastic collision process $i$, respectively. $T_e$ is electron temperature. $\vec{E}$ and $\rho$ are electric field and total charge density, respectively. $n_\alpha(t = -\infty)$ is the initial density of charged species $\alpha$. $n_\alpha, \bar{u}_\alpha$ are the density and fluid velocity of charged species $\alpha$. $N_O, N_{O2}, N_{N2}$ are the densities of the three neutral species. (b) Coordinate system $(r, z)$ used in simulation. Maximum radius is 64 km; maximum altitude is 90 km.
<table>
<thead>
<tr>
<th>Species</th>
<th>Altitude (km)</th>
<th>Collision Frequency</th>
<th>Cyclotron Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>ions</td>
<td>60</td>
<td>93 kHz</td>
<td>27 Hz</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>20 kHz</td>
<td>27 Hz</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>760 Hz</td>
<td>27 Hz</td>
</tr>
<tr>
<td>electrons</td>
<td>60</td>
<td>22 MHz</td>
<td>1.5 MHz</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>4.8 MHz</td>
<td>1.5 MHz</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>180 kHz</td>
<td>1.5 MHz</td>
</tr>
</tbody>
</table>

Table 2.1: Momentum transfer collision frequencies of electrons and ions with neutrals, and cyclotron frequencies of electrons and ions at representative altitudes. The source of these collision frequencies is explained in Section 2.5.

To quantitatively justify treating the medium as unmagnetized, we examine the mobility tensor for particles in a z-directed magnetic field [Krall and Trivelpiece, 1986, p. 329]:

\[
\mu = \begin{bmatrix}
\mu_\perp & \mu_T & 0 \\
\mu_T & \mu_\perp & 0 \\
0 & 0 & \mu_{\|}
\end{bmatrix}
\]  

(2.1)

where

\[
\mu_\perp = \frac{\mu_0 v_m^2}{v_m^2 + f_c^2}; \quad \mu_T = \frac{\mu_0 f_c v_m}{v_m^2 + f_c^2}; \quad \mu_{\|} = \mu_0 = \frac{q}{m v_m}
\]  

(2.2)

and \(q\) is particle charge. As we will see in Section 2.7, the mobility tensor specifies the steady state particle dynamics in the presence of an electric field. Particle drift velocity is expressed in terms of the mobility as:

\[
\vec{u} = \mu \cdot \vec{E}
\]  

(2.3)

In the case of collision frequency greatly exceeding cyclotron frequency, \(v_m >> f_c\), we find

\[
\mu_\perp \equiv \mu_0 \quad \text{and} \quad \mu_T << \mu_0
\]  

(2.4)

so that the mobility tensor becomes diagonal, and the mobility is isotropic and equal to the scalar \(\mu_0\). As anticipated, this is precisely the statement that the background magnetic field does not influence particle dynamics.

One might note from the last entry in Table 2.1 that the electron cyclotron frequency exceeds the electron-neutral collision frequency in a thin layer at the top of the simulation volume. However, we will see that because of the extremely high conductivity at these altitudes, the electric field vanishes there, and the electrons are essentially stationary. As a result, the magnetic field still does not affect particle dynamics.
We will now discuss each of the component models that make up the comprehensive three-dimensional model. We will start with the supporting models and conclude with the three-fluid model itself.

2.2 Modeling the Ambient Medium

The charged and neutral components of the ambient medium are treated separately. Semi-empirical models are employed in both cases. Data are taken for the atmosphere and lower ionosphere above the point on the earth at $35.1^\circ$ N latitude, $260.4^\circ$ E longitude, which is the approximate location of Texola, Oklahoma, USA. This location on the border between Texas and Oklahoma was chosen because it lies amid the lower plains states—a region notable for its frequent and severe summertime thunderstorms—and because sprites were actually observed there. The region averages nearly fifty thunderstorm days per year [Uman, 1987, p. 42]. During the Sprites94 measurement campaign, several sprites were observed almost directly over Texola [Sentman and Wescott, 1995].

Ambient conditions were chosen to represent 2200 hr local time on 4 July. Thunderstorms are known to occur on this date and time, and the after-dusk time of day would be favorable for the observation of sprites [Lyons, 1994].

2.2.1 The Neutral Atmosphere

Neutral temperature, density, pressure, and composition in the altitude range 0 - 100 km are obtained from the COSPAR International Reference Atmosphere (CIRA) [COSPAR, 1972] for July at $30^\circ$ N latitude. The initial neutral properties are assumed to be uniform in the horizontal dimensions over the 120 km diameter of the simulation volume. Values at an arbitrary altitude are gotten by cubic polynomial interpolation on the CIRA points. The necessary coefficients are obtained using Neville’s algorithm [Press et al., 1994]. The interpolated fit and the CIRA values for neutral density, temperature, and pressure are shown in Figure 2.2.

The neutral atmosphere is assumed to be motionless, so that the only variation in neutral density results from impact ionization by heated electrons. This process removes particles from the neutral population and places them in the ion population. Due to the exceedingly small relative density of electrons (see Section 2.2.2) the resulting fractional change in neutral density will turn out to be infinitesimal, as discussed in Section 2.5 regarding collisional processes. Neutral density
Figure 2.2: Model of neutral atmosphere employed in this work, obtained from COSPAR International Reference Atmosphere (CIRA). Circles represent CIRA points, solid lines are interpolations. (a) Neutral temperature, T, and pressure, P. (b) Particle densities of the three dominant neutral components, atomic oxygen, molecular oxygen, and molecular nitrogen.
will therefore be considered constant. We will ignore heat transfer to the neutral populations and justify *a posteriori* the assumption that it is negligible. The neutral temperature is therefore considered fixed.

The CIRA model for atmospheric neutral temperature and density thus provide the initial values for the densities of the neutral species and the constant values of neutral temperature for use by the rest of the simulation.

2.2.2 The Charged Constituents

The quiescent ionic and electronic content of the atmosphere is established by several processes that are considered in detail in the chemical effects model of Section 2.6. At altitudes up to 50 km, the principal source of ionization is cosmic radiation [Sagalyn and Burke, 1985]. Within the lowest 2.5 km (known as the planetary boundary layer), over land surfaces, cosmic radiation is strongly supplemented by radiation from radioactive substances in the soil. Above 50 km, the largest daytime source of ions is the photo-ionization of nitric oxide (NO) by solar radiation. Chemical reactions among the electrons, neutrals, and ionic constituents of the mesosphere and the ionospheric D region lead to the production of the principal small positive ions NO⁺ and O₂⁺, the principal negative ions NO₃⁻ and O₃⁻, as well as a variety of large positive ions known as cluster ions [Mitra, 1981]. These reactions will be considered at length in Section 2.6.1.

The several positive and negative ion species mentioned above are treated in the aggregate as either positive or negative ions. The small ion designation is made in distinction to the large aerosol particles and hydrated ions that exist in measurable quantities at altitudes ranging from the earth’s surface all the way up to the mesosphere. The larger ionic species are considerably more massive than the small ions and have correspondingly lower mobility. One representative large ion, H⁺•(H₂O)₅, has a mass of 91.09 amu compared to NO⁺, with a mass of 30.01 amu [Adler, 1993, p. 168]. This greater mass results in a significant reduction in large ion mobility below even that of the small positive ions [Reid, 1986, p.189]. The large ions are thus essentially immobile and do not contribute significantly to the total conductivity of the medium.

The quiescent values of small positive ion density used as an initial condition in the simulation are obtained from balloon soundings of the upper atmosphere taken over the American Midwest [Rosen and Hofmann, 1981a]. Multiple soundings have shown that the mean positive polar conductivity, σ⁺, in units of 10⁻¹⁴ per ohm-meter, is related to the neutral pressure, P, in millibars, by the smooth curve [Rosen and Hofmann, 1981b]:

41
\[ \sigma_+ = 3789 \cdot P^{-0.6975} \cdot e^{-0.002899 P} \]  \hspace{1cm} (2.5)

This holds, independent of altitude, at all altitudes from the ground up to at least 60 km. Positive polar conductivity is the contribution of the small positive ions to the total conductivity, namely:

\[ \sigma_+ = n_p e \mu_s \]  \hspace{1cm} (2.6)

where \( n_p \) is small positive ion density, \( e \) is the electronic charge, and \( \mu_s \) is a generic small ion mobility. Positive and negative small ions differ in mobility by less than roughly 10% [Gringel et al., 1986], so we do not track separate mobilities for the two varieties. Furthermore, it is found that the reduced small ion mobility, \( \mu_R \), defined as the mobility normalized to standard temperature and pressure \((P_0 = 1013\) millbar, and \( T_0 = 273^\circ \) K), is roughly independent of altitude [Rosen and Hofmann, 1981b]:

\[ \mu_R = \frac{T_0 P}{T P_0} \cdot \mu_s = 1.5E-04 \text{ m}^2/(\text{Vs}) \]  \hspace{1cm} (2.7)

where \( T \) is ambient neutral temperature. We can now combine (2.5), (2.6), and (2.7) to obtain small positive ion density as a function of altitude, \( z \).

\[ n_p(z) = \frac{1}{e \mu_R} \cdot \frac{T_0 P(z)}{T(z) P_0} \cdot 3789 \cdot 10^{-14} \cdot P(z)^{-0.6975} \cdot e^{-0.002899 P(z)} \text{ (per m}^3) \]  \hspace{1cm} (2.8)

This density profile is accurate at altitudes below the onset of significant electron density at 70 km. Additional information is required to extend this model for \( n_p(z) \) to higher altitudes, as discussed below. By rearranging (2.7), we find small ion mobility as a function of altitude:

\[ \mu_s(z) = \frac{T(z) P_0}{T_0 P(z)} \cdot \mu_R \]  \hspace{1cm} (2.9)

The initial electron density profile employed in the simulation is a composite of two modeled results for the altitude ranges above and below the sharp ledge in electron density near 75 km. Electron density above 75 km is taken from the International Reference Ionosphere (IRI-90) [Bilitza, 1990]. Below the sharp ledge near 75 km, electron density falls rapidly with decreasing altitude and exhibits considerable variability in time. The range of possible profiles is suggested by Rowe et al. [1974]. A typical profile shows an ambient electron density of 5.5E+10 per m\(^3\) at 80 km, consistent with the IRI-90 model for 10 p.m. local time on a July night in Oklahoma under conditions of moderate solar activity \((R_{Z12} = 50)\). Below this, electron density falls with a hypothetical decade scale height of between 2.5 km and 7.5 km, depending upon the conditions under study. The case of a 2.5 km scale height corresponds to a tenuous nighttime D layer, while
the 7.5 km scale height corresponds to a dense layer. The proposed range of profiles is consistent with the range of observed electron density profiles in this altitude regime presented by Rowe et al. [1974]. A 7.5 km decade scale height for nighttime electron density in the altitude range 50 to 70 km was reported by Swider et al. [1978].

Under quiescent conditions, the negative and positive ion populations exhibit equal densities at all altitudes from the earth up to the altitude at which the free electron population becomes appreciable. At altitudes above 60 km, electrons begin to take over from small negative ions as the dominant negative charge carrier. At altitudes above 70 km, electrons are vastly more numerous than small negative ions, and positive ion density is precisely balanced by electron density. Throughout the transition from small negative ions to electrons, charge neutrality is maintained. In fact the only departure from strict charge neutrality in the quiescent system is a layer of space charge, $\rho_s$, near the ground that accounts for the small fair weather electric field discussed in Section 2.4.1.

The small positive ion density profile of (2.8) is accurate at altitudes where small negative ions are the majority negative charge carrier, that is, below 70 km. Going upward through the 70 km transition, the relative electron and ion densities rapidly change from $n_e \ll n_p$ to $n_e \gg n_p$, where $n_p$ at high altitude is obtained by extrapolating the low altitude result (2.8). Below the transition, quasi-neutrality is maintained under a balance between small negative and positive ion density. Above the transition, a physically unrealistic preponderance of electrons over small positive ions arises if we insist on extrapolating (2.8) into the region. Instead, we will use (2.8) as a first guess at small positive ion density at all altitudes:

$$a(z) = n_p(z)$$  \hspace{1cm} (2.10)

We then obtain the correct high altitude, small positive ion profile by adding in the modeled electron density, $n_e(z)$.

$$n_p(z) = a(z) + n_e(z)$$  \hspace{1cm} (2.11)

At altitudes below the negative carrier transition, the electron density has a negligible effect on the prediction of (2.8). At higher altitudes, (2.11) assures that the small positive ion density will equal the ambient electron density. The altitude profile of the small positive ion density according to this model are presented in Figure 2.3.

The small positive ion density profile obtained in (2.11) is consistent both with the observations of Rosen and Hofmann [1981a] and with the principle of steady state charge neutrality. However, while physically justifiable, (2.11) was obtained without reference to the
Figure 2.3: Small positive ion density for use as an initial condition in the simulation. $n_e$ and $n_p$ denote curves of electron and small positive ion density, respectively. Dashed curve of modeled electron density is added in to extend low altitude positive ion density model to altitudes above 70 km.
small negative and cluster ion density. In fact, we have to this point not attempted to model the
electron, small negative ion, and cluster ion densities in a mutually consistent manner. To arrive at
mutually consistent density profiles, we will treat the small positive ion density (2.11) as ground
truth, and determine the densities of the remaining charged particles to be consistent with it—
according to the steady state chemical balance among all of the ions and electrons (see Section
2.6.2). Built into the resulting equilibrium is steady state charge neutrality.

In this manner, the crossover from small negative ions to electrons as the high altitude
negative charge carrier is addressed, while steady state charge neutrality is maintained. With all of
these considerations satisfied, the initial positive ion density profile employed in the simulation
obeys:

\[ n_p(z) = \frac{1}{e} \rho_f(z) + n_e(z) + n_m(z) - n_e(z) \]  \hspace{1cm} (2.12)

where \( n_e(z) \) is the cluster ion density, and \( \rho_f(z) \) is the charge density associated with fair weather
electric field, derived in Section 2.4.1. This ensures that the fair weather electric field is consistent
with the low altitude difference between the positive and negative ion densities. At the highest
altitudes, the first and third terms vanish, and electrons precisely balance the positive ions.

The total ambient conductivity \( \sigma(z) \) is a function of the mobility and number density of
each of the charged species:

\[ \sigma(z) = \sum_{\alpha} q_\alpha n_\alpha(z) \mu_\alpha(z) \]  \hspace{1cm} (2.13)

where \( q_\alpha \) is the charge on species \( \alpha \). The electron and ion densities are specified in the initial state
by the considerations outlined above. Their subsequent evolution is determined from the fluid
model developed in Section 2.7. The small ion conductivity is given by (2.9). It remains only to
determine the electron mobility. In an unmagnetized plasma, the isotropic electron mobility is
given by (2.2):

\[ \mu_e = -\frac{e}{m_e \nu_m} \]  \hspace{1cm} (2.14)

where \( m_e \) is the electron mass and \( \nu_m \) is the momentum transfer collision frequency for electron-
neutral collisions. In Section 2.5 we will discuss the manner in which we obtain the momentum
transfer collision frequency from the energy-dependent, electron-neutral scattering cross sections.
2.3 Cloud Charge and Lightning Model

The cloud charge model addresses the spatial distribution of charge density used to represent the real world situation described in Section 1.1.2 on lightning phenomenology. Since the present formulation is not intended to treat short-timescale electromagnetic effects, the direct influence of the rapidly changing lightning channel current is not considered. As a result, the lightning model simply addresses the time evolution of the cloud charge density, and both lightning and cloud charge are specified by a unified description.

The introduction and removal of cloud charge are treated as external stimuli to the system, and the purpose of the present simulation is to predict the response of the system to those stimuli. The physical processes responsible for cloud electrification and subsequent lightning discharge are beyond the scope of this effort. However, we may use the sizable body of observations regarding these phenomena to form an empirical model to characterize them.

The spatial distribution of cloud charge is described by two, three-dimensional Gaussian forms, one with positive amplitude and centered at an altitude \( a \), representing the positive charge center of the cloud, and one with negative amplitude at an altitude \( a_n \), representing the negative charge center. The two charge centers are aligned vertically, with \( a > a_n \). The radial scale size of each Gaussian is denoted by \( b \), roughly representing the radius of the charge blob. The time of peak cloud charge density occurs just at the initiation of the lightning discharge at time \( t = 0 \). At time \( t = 0 \), the cloud source charge distribution would appear:

\[
\rho_s(\vec{r}) = \frac{Q_0}{\pi^{3/2}b^3} \exp \left( - \frac{\vec{r} - a\hat{z}^2}{b^2} \right) - \frac{Q_0}{\pi^{3/2}b^3} \exp \left( - \frac{\vec{r} - a_n\hat{z}^2}{b^2} \right)
\]  

(2.15)

where \( \hat{z} \) is a unit vector in the vertically upward direction. Each term in the distribution is normalized in such a manner that the total charge associated with each center is \( Q_0 \). Typical values for the parameters involved are \( Q_0 = 100 \, \text{C} \); \( b = 3 \, \text{km} \); \( a = 15 \, \text{km} \); \( a_n = 8 \, \text{km} \). A pair of surfaces of constant cloud charge density consistent with these parameters is illustrated in Figure 2.4a.

We represent cloud electrification and subsequent lightning discharge by the proportionate growth and diminution of charge density everywhere within the distribution. We accomplish this by imposing a time dependence on the coefficients \( Q_0 \). The positive and negative charge centers each undergo a separate time evolution in the dipole cloud charge model of (2.15). During the electrification period, the total charge in the positive and negative sections is maintained as
Figure 2.4: Spatial and temporal characteristics of cloud charge and lightning model. (a) Spherical surfaces of constant cloud charge density of radius 3 km (representing contours of roughly 10% of peak density for the bipolar cloud charge model). Dimensions of plot are chosen to suggest the relative size and separation of the cloud charge centers versus the simulation volume. Band in xy-plane represents footprint of cylindrical simulation boundary. (b) Cloud charge evolution showing electrification and discharge periods on the same time scale. Lightning discharge initiates at time = 0. Broken curve shows quadratic growth in time for a 20 s electrification period. Solid curve shows exponential growth for the case of a typical 1.25 s time constant. (c) Evolution of cloud charge during discharge and tail of electrification period. Quadratic and exponential curves overlay each other.
equal and opposite—representing the transfer of charge between the two. During the discharge period of a positive CG discharge, the density of the positive charge distribution is reduced while the negative charge distribution is held constant. Similarly, during the discharge period of a negative CG discharge, the lower, negative charge center is diminished, while the upper, positive charge center is held constant. The behavior of the total charge, $Q_+$ and $Q_-$, in the two charge centers is represented schematically in Figures 2.5a and 2.5b for the cases of positive and negative CG discharges in the dipole model. The bottom line is that we must replace (2.15) by the time-dependent version:

$$\rho_S(\vec{r}, t) = \frac{Q_+(t)}{\pi^{3/2}b^3} \exp\left(-\frac{|\vec{r} - \vec{a}_+|^2}{b^2}\right) - \frac{Q_-(t)}{\pi^{3/2}b^3} \exp\left(-\frac{|\vec{r} - \vec{a}_-|^2}{b^2}\right)$$

(2.16)

The precise forms of $Q_+(t)$ and $Q_-(t)$ must be given considerable latitude for variation.

Both electrification and discharge have been observed to take place over a range of time scales (see Section 1.1.2), and electrification exhibits considerable variation in the shape of the cloud charge waveform. In particular, cloud electrification can take place over periods of tens of seconds to tens of minutes [Israel, 1973, p. 523; Smith, 1958; and Levin and Tzur, 1986, p. 137]. The longer electrification periods tend to be roughly quadratic in time, while the shortest ones are nearly exponential. These options during the electrification period are illustrated in Figure 2.4b. The discharge waveforms, on the other hand, are generally exponential. A typical case is illustrated in Figure 2.4c. Note that in our model, the electrification always takes place during the interval for which $t < 0$; the lightning discharge initiates at precisely $t = 0$; and the discharge and relaxation unfold while $t > 0$.

Up to this point, we have been describing the electrification and discharge processes in terms of the dipole model of cloud charge structure, that is, a charge structure consisting of vertically aligned centers of positive and negative charge, with the positive charge located at higher altitude than the negative. We noted in Section 1.1.2, however, that in cases of positive cloud-to-ground lightning, the positive charge center can be displaced horizontally from the negative by as much as tens of kilometers. In such cases—and these are the cases of primary interest when considering sprites—it may be more appropriate to consider modeling the cloud charge as an isolated positive monopole at altitude $a$. In the monopole model there is no negative cloud charge center. Cloud charge evolution in the monopole model is illustrated in Figure 2.5c.
Figure 2.5: Behavior of cloud charge under monopole and dipole models. Time scale of discharge is unrealistically long for purposes of illustration. (a) Positive cloud-to-ground (CG) discharge under dipole cloud charge model. (b) Negative CG discharge under dipole model. (c) Positive CG discharge under monopole model.
In the interest of preserving computational efficiency, we restrict ourselves to cylindrically symmetric charge configurations. This leads us to consider only the monopole and the vertically aligned dipole models. Reality is in between the two, perhaps closer to the monopole configuration.

2.4 Electric Field Model

Our most fundamental assumption regarding the electric field—indeed, regarding our whole approach to lightning effects—is that we may consider it to be quasi-static. This approach is justified by the fact that we are interested in physical effects that evolve over many milliseconds. For comparison, the transit time for a propagating electromagnetic disturbance over the longest dimension of the system (90 km) is only 300 µs. This places us in the low frequency, long wavelength limit of Maxwell’s Equations. As a result, we can ignore the time derivative of magnetic flux density in Faraday’s Law:

\[ \nabla \times \vec{E} = -\mu \frac{\partial \vec{H}}{\partial t} \]  \hspace{1cm} (2.17)

We estimate the orders of magnitude of the left and right hand sides:

\[ \text{LHS} \sim \frac{\eta H_0}{\Delta l} = \sqrt{\frac{\mu}{\varepsilon} \frac{H_0}{\Delta l}} \quad \text{and} \quad \text{RHS} \sim \frac{\mu H_0}{\Delta t} \]  \hspace{1cm} (2.18)

where \( \eta \) is the wave impedance of the medium, and \( \Delta l \) and \( \Delta t \) are the length and time scales of the system, respectively. In the present case, for which \( c \Delta t \gg \Delta l \), we conclude that \( \text{LHS} \gg \text{RHS} \), so:

\[ \nabla \times \vec{E} = 0 \Rightarrow \vec{E} = -\nabla \varphi \]  \hspace{1cm} (2.19)

and the electric field can be derived from the electrostatic potential, \( \varphi \).

At low altitudes (below roughly 60 km), finding the electric field is less a matter of physical modeling than of numerical techniques. In this altitude regime, charge density is determined by continuity of current, and the electrostatic field is obtained by a straightforward (in principle) application of Poisson’s Equation:

\[ \nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \]  \hspace{1cm} (2.20)
where $\rho$ accounts for all of the free charge in the plasma medium, and there is no bound charge, so that it is appropriate to use the permittivity of vacuum, $\varepsilon_0$, as the uniform permittivity of the medium. Any difficulties with this approach are principally numerical and will be discussed in Chapter 3. The only physical considerations to be addressed here are the boundary conditions.

The highest reaches of the simulation volume lie in the lower ionosphere, where the charge density is everywhere very nearly zero. This fact presents difficulties for a solution of (2.20), in that numerical inaccuracies can overwhelm the tiny amount of charge actually present. We will therefore develop an alternative means of computing the high-altitude electric field that is both numerically tractable and physically more appropriate.

In the coming sections, we consider first the quiescent atmospheric electric field, then the boundary conditions to be imposed on the electric field solution, and finally the details of the high altitude electric field model.

2.4.1 Fair Weather Electric Field

Under the quiescent conditions that exist away from storms, a steady downward electric field and associated current are always observed in the atmosphere. These are known as the fair weather electric field and current. The fair weather electric field observed at the ground is directed vertically downward and exhibits a typical magnitude between 100 and 150 V/m. The fair weather current is one part of a global electrical circuit that comprises upward current through storms, lateral current through the ionosphere, downward current through fair weather regions and lateral current once again through the earth back to the storms [Roble and Tzur, 1986]. Significant input to the global electrical circuit also comes from sea spray and volcanic eruptions. The total global current has been estimated at 1 - 2 kA. Spread over the surface of the earth, this amounts to a current density on the order of $10^{-12}$ A/m$^2$. A disproportionately large amount of current flows to mountains, where the columnar resistance of the atmosphere is reduced.

The fair weather electric field constitutes the initial condition for the electric field in the simulation, $\vec{E}(r, z, t = -t_e)$, where $t_e$ is the duration of the electrification period, and lightning initiates at $t = 0$. To determine the fair weather electric field as a function of altitude, we postulate a value of the field at the earth's surface, $-E_0$, and use it as a boundary condition along with the altitude profile of ambient total conductivity. Under the assumption of a spherically symmetric
atmosphere, the fair weather electric field is a function of altitude only, \( E_f(z) \), and is directed along the z-axis:

\[
\bar{E}(r, z, t = -t_e) = \hat{z}E_f(z)
\]  

(2.21)

Continuity of current in the steady state implies that the current density is divergence-free, which, in the case of earth-centered, spherical symmetry, we express as:

\[
\frac{1}{R^2} \frac{d}{dR} \left( R^2 J_R \right) = 0
\]  

(2.22)

where \( R \) is the radial coordinate from the earth’s center, and \( J_R \) is the radial component of the current density in the earth-centered system. \( J_R \) is equivalent to the vertical component of current density in the simulation coordinates, in this instance the fair weather current density, \( J_f \). We integrate (2.22) to obtain:

\[
J_R(R) = J_f(R) = \frac{J_0 R_E^2}{R^2}
\]  

(2.23)

where \( J_0 \) is the fair weather current density at the earth’s surface, \( R_E \) is the earth’s radius. We use \( R = R_E + z \) to write:

\[
J_f(z) = -\frac{\sigma_0 E_f R_E^2}{(R_E + z)^2}
\]  

(2.24)

where \( \sigma_0 \) is the atmospheric conductivity at the earth’s surface. Finally, we apply Ohm’s Law (2.3) for the case of isotropic ambient conductivity, \( \sigma(z) \), to conclude:

\[
E_f(z) = -\frac{\sigma_0 E_f R_E^2}{(R_E + z)^2 \sigma(z)}
\]  

(2.25)

The fair weather electric field obtained from this model, along with the total ambient conductivity, is plotted in Figure 2.6.

The fair weather charge density associated with this field is obtained by application of Poisson’s Equation in spherical coordinates:

\[
\rho_f(R) = \frac{1}{R^2} \frac{d}{dR} \left( R^2 \varepsilon_0 E_f \right)
\]  

(2.26)

where we have used the permittivity of free space because the fair weather electric field is appreciable only in the bottom twenty kilometers of the neutral atmosphere. In terms of the altitude, \( z \), this evaluates to:

52
Figure 2.6: Fair weather electrical properties of the atmosphere. (a) Fair weather electric field profile, $E_f$, for surface level field strength of 100 V/m. Total conductivity, $\sigma$, of quiescent medium. (b) Charge density associated with fair weather electric field.
\[ \rho_f(z) = -\epsilon_0 \sigma_0 E_0 R_E^2 \frac{d}{dz} \frac{1}{(R_E + z)} \sigma(z) \]  \hspace{1cm} (2.27)

2.4.2 Boundary Conditions

Despite the fact that the high-altitude solution of Poisson’s Equation is subject to slight but significant inaccuracy due to the extremely small values of high-altitude charge density, we nevertheless solve (2.20) throughout the simulation volume. We subsequently replace the questionable, high-altitude portion of the solution with an alternative solution described in Section 2.4.3. Thus our first step is to solve Poisson’s Equation subject to conditions imposed at the boundaries of the entire simulation volume.

The spatial boundaries of the electrostatic field solution are a cylindrical surface of radius 64 km that is coaxial with the 2-axis, the flat bottom of the cylinder at the earth surface, and the flat top at 90 km altitude. We will examine the conditions at each of the boundaries.

The bottom boundary is the air-soil interface. At very low frequencies, the rich, damp soils that one might expect to find in Oklahoma during a thunderstorm have a conductivity, \( \sigma \), of 1.0E-02 to 3.0E-02 per ohm-meter and a relative permittivity, \( \epsilon_r / \epsilon_0 \), of 15 to 30 [Watt, 1967, p. 183]. For these values and a representative angular frequency, \( \omega \), of 2\( \pi \times 10^3 \) rad per s, the so-called loss tangent, \( \sigma / (\omega \epsilon_1) \), becomes roughly 2.0E+04. The choice of angular frequency was made to be consistent with our interest in phenomena that evolve on time scales of 1 millisecond or longer. The fact that the loss tangent is so large compared to unity tells us that the medium in question is a nearly perfect conductor [Watt, 1967, p. 185]. We may explain this conclusion by noting that the coefficient of reflection for a transverse electric plane wave incident from air (medium 0) on medium 1 is:

\[ R = \frac{1 - p_{01}}{1 + p_{01}} \text{ where } p_{01} = \frac{\mu_0 k_{1z}}{\mu_1 k_{0z}} \] \hspace{1cm} (2.28)

and \( k_z \) is the component of the wavenumber normal to the boundary in each medium [Kong, 1990, p. 119]. The wavenumber in air is just the free space wavenumber, while the wavenumber in the conductive medium 1 is determined from the dispersion relation [Zahn, 1987, p. 509]:

\[ k_{1z}^2 = \omega^2 \mu_0 \epsilon_1 \left( 1 + \frac{\sigma}{j\omega \epsilon_1} \right) \] \hspace{1cm} (2.29)
where \( j \) is the imaginary unit, and the wave fields have a time dependence of the form \( e^{j \omega t} \). Using this result along with the fact that the permeability of both media, \( \mu_0 \) and \( \mu_1 \), are that of free space, we find:

\[
P_{01}^2 = \frac{\varepsilon_1}{\varepsilon_0} \left( 1 + \frac{\sigma}{j \omega \varepsilon_1} \right)
\]  
(2.30)

For the value of the loss tangent found above, the reflection coefficient given by (2.28), in conjunction with (2.30), approaches minus one. A reflection coefficient of \( R = -1 \) corresponds to the condition of vanishing tangential electric field at the air-soil boundary, since in that case the incident and reflected fields precisely cancel at the boundary. This corresponds to an interface between air and a perfect conductor, above which tangential electric field must be zero.

We conclude from this analysis that the tangential electric field at the earth boundary of the simulation volume is zero to an excellent approximation. We have considered electromagnetic waves in validating this conclusion in spite of the fact that the simulation is electrostatic. This was done to ensure that the conclusion remains valid for the slowly changing fields we expect to encounter. Moreover, in the static limit, we see from (2.30) and (2.28) that

\[
\lim_{\omega \rightarrow 0} R = -1
\]  
(2.31)

and the approximation of zero tangential field becomes exact.

We translate the conductive boundary condition into the following specification for the electrostatic potential on the earth boundary of the simulation volume:

\[
\varphi(r, z = 0) = 0
\]  
(2.32)

The potential is constant and is arbitrarily assigned the value 0.

In a similar vein, we note from Figure 2.6 that the total conductivity at 90 km altitude is again on the order of \( 10^2 \) per ohm-meter, just as for the earth surface. As a result, the upper boundary of the simulation volume approximates an equipotential surface every bit as well as the earth boundary, and we write:

\[
\varphi(r, z = z_{\text{max}}) = V_0
\]  
(2.33)

where \( V_0 \) is known as the ionospheric potential, and \( z_{\text{max}} \) is the maximum altitude in the simulation volume, 90 km. We consider the ionosphere to be a great spherical electrode, whose potential is fixed by global processes external to our simulation system. Some 1500-2000 thunderstorms are active somewhere on the planet at any given time [Roble and Tzur, 1986]. Taken together, these storms exercise decisive influence in determining the ionospheric potential, but the effect of any
single storm is negligible. We therefore consider $V_0$ to be a parameter that is fixed in time for the duration of the simulation. Measurements show a mean value of 280 kV [Roble and Tzur, 1986], but for the sake of the self-consistency of the simulation we fix $V_0$ at the initial value given by:

$$V_0 = - \int_0^{z_{inf}} E_f(z) \, dz$$  \hspace{1cm} (2.34)$$

which, for the fair weather electric field given by (2.25), yields a value of 320 kV.

The correct boundary condition at the cylindrical surface is less clear cut. We will implement two equally valid conditions and verify that the choice of one or the other does not significantly affect the electric field solution. We can specify either the electrostatic potential or its radial derivative (equivalently, the radial electric field) at all points on the cylindrical surface. Unfortunately, we do not have accurate a priori knowledge of either quantity. The best we can do is to assume that because the radius of the simulation volume, 64 km, is much greater than the radial scale size of the cloud charge distribution, 3 km, ambient conditions will exist everywhere on the cylindrical surface. Ambient conditions amount to:

$$\varphi(r = a, z) = - \int_0^{z} E_f(\zeta) \, d\zeta$$  \hspace{1cm} (2.35)$$

$$\frac{\partial}{\partial r} \varphi(a, z) = 0$$  \hspace{1cm} (2.36)$$

We may apply either (2.35) or (2.36) as the boundary specification, knowing that both are approximately correct for a simulation volume of sufficient radial dimension.

In order to assess the reasonableness of this approach, we can check that in practice one condition roughly implies the other. The results of such a test are displayed in Figure 2.7. Figure 2.7a shows the results of an electric field calculation using the boundary condition expressed in (2.35). We plot the radial electric field at the edge of the simulation volume, normalized to the maximum radial field in the volume. The radial electric field at the edge is less than 0.2% of maximum, which is very small—consistent with (2.36). Figure 2.7b shows the electrostatic potential at the edge of the simulation volume, calculated for the boundary condition of (2.36). Plotted for comparison is the fair weather electrostatic potential. While the two curves are not identical, they are in general agreement.

As a result of these findings, we suggest that the two boundary conditions expressed in (2.35) and (2.36) produce equivalent, and reasonable, effects on the electric field calculation.
Figure 2.7: Measures of effectiveness of boundary condition approximations. Taken from simulation results for 50 C peak cloud charge. (a) Radial electric field at cylindrical boundary surface, normalized to maximum radial field in simulation volume. Calculated for electrostatic potential specified at quiescent values on cylindrical surface. Shows that boundary condition (2.35) leads to negligible radial electric field (< 0.2% peak) at cylindrical boundary. (b) Comparison of fair weather electrostatic potential profile (broken line) and calculated potential for boundary condition of zero radial electric field at cylindrical surface (solid line). Shows that boundary condition (2.36) leads to reasonable potentials on cylindrical surface of simulation volume.
Either condition is an acceptable approximation, particularly since we are not especially interested
in the solution near the cylindrical simulation boundary. For philosophical compatibility with the
high altitude solution presented in Section 2.4.3, we will apply the boundary condition (2.36),
namely that the radial electric field is zero at the cylindrical boundary.

2.4.3 High Altitude Electric Field Solution

We have pointed out that above some altitude, typically near 70 km, the charge density is
especially zero, and we can accurately approximate the medium as neutral. In the quasi-neutral
medium we apply a technique to solve for the electric field that differs from the one used in non-
neutral regions. This not only improves numerical accuracy and tractability, but also ensures that
the means of determining the electric field is physically justifiable.

Regardless of the nature of the medium, the electric field, charge density and current
density must simultaneously satisfy three physical principles in the steady state: the Poisson
Equation (2.37), continuity of current (2.38), and Ohm's Law (2.39).

\[ \nabla \cdot (\varepsilon \vec{E}) = \rho \quad (2.37) \]

\[ \frac{\partial}{\partial t} \rho + \nabla \cdot \vec{J} = 0 \quad (2.38) \]

\[ \vec{J} = \sigma \vec{E} \quad (2.39) \]

This expression of continuity of current comes from a fluid formulation that will be fully discussed
in Section 2.7. Likewise, the use of Ohm's Law to summarize charged particle dynamics in the
drift limit is justified in Section 2.7.

In the case of the non-neutral medium, the electric field at a given point in time determines
the current density according to Ohm's Law. Based upon this current density, continuity of current
then determines the evolution of the charge density to a point in time slightly in the future. The
electric field at this future time is then calculated from Poisson's Equation. Repeated cycles of this
description step the solution into the future.

The same equations must be satisfied in the case of the quasi-neutral medium. However,
because the charge density is exceedingly small, it is not only inaccurate, but physically
inappropriate to determine the electric field from the charge density using (2.37). Instead, we solve
current continuity in the limit of negligible charge density:
\[ \vec{V} \cdot \vec{J} = 0 \]  
(2.40)

We can combine this directly with Ohm’s Law to obtain a static equation for the electric field:

\[ \vec{V} \cdot (\sigma \vec{E}) = 0 \]  
(2.41)

We thus determine the electric field directly from current continuity, and use Poisson's Equation to calculate the charge density. Just as in the non-neutral case, (2.37) through (2.39) are satisfied. However, we abandon the Poisson Equation as a means of calculating the electric field at high altitudes because the plasma there is known to be quasi-neutral. We thereby prevent physically insignificant amounts of charge from determining the macroscopic electric field. Instead, we determine the field from macroscopic dynamics, as expressed by Ohm’s Law, together with the approximation that charge density is zero. We validate this approximation by subsequently calculating charge density from the electric field using the Poisson Equation. The charge density determined in this manner arises as a consequence of variations in conductivity in the presence of approximately uniform current density—which is the physically appropriate mechanism.

We proceed with the high altitude electric field solution by rewriting (2.41) as:

\[ \vec{V} \cdot \vec{E} + \vec{r} \cdot \vec{E} = 0, \text{ where } \vec{r} = \frac{\vec{V} \sigma}{\sigma} \]  
(2.42)

which, in terms of the electrostatic potential, becomes:

\[ \nabla^2 \varphi + \vec{r} \cdot \nabla \varphi = 0 \]  
(2.43)

In axisymmetric cylindrical coordinates, this takes the form:

\[ \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \varphi}{\partial r} \right) + \frac{\partial^2 \varphi}{\partial z^2} + \frac{\partial \varphi}{\partial r} + \Gamma \frac{\partial \varphi}{\partial r} + \Gamma z \frac{\partial \varphi}{\partial z} = 0 \]  
(2.44)

We seek to simplify this equation by proposing that

\[ \left| \Gamma z \frac{\partial \varphi}{\partial z} \right| \gg \left| \Gamma r \frac{\partial \varphi}{\partial r} \right| \]  
(2.45)

This condition is unequivocally justified in the quiescent state, since then both the field and the conductivity gradient are purely in the z-direction. Moreover, it is plausible that even under conditions of thundercloud electrification and discharge, the dominant electric field component at high altitudes will be vertical, directed along a line from the cloud charge. The simulation results of Chapter 4 will bear this out. Furthermore, we will see that lightning-induced radial conductivity variations at high altitude exhibit factor-of-two scale lengths on the order of 10 km, whereas the
corresponding vertical scale length is less than a kilometer due to the extremely rapid vertical variation in neutral density (see Figure 2.6a).

The overwhelming rapidity of the conductivity variation in the $z$-direction also permits us to take $\Gamma_z$ to be a function of $z$ only. Applying this fact along with (2.45), we find that (2.44) reduces to:

$$\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \varphi}{\partial r} \right) + \frac{\partial^2 \varphi}{\partial z^2} + \Gamma_z(z) \frac{\partial \varphi}{\partial z} = 0 \quad (2.46)$$

which can be solved by separation of variables using a solution of the form:

$$\varphi(r, z) = A(r)B(z) \quad (2.47)$$

We obtain two equations by substituting (2.47) in (2.46).

$$\frac{1}{r} \frac{d}{dr} \left( r \frac{dA}{dr} \right) + k^2 A(r) = 0 \quad (2.48)$$

$$\frac{d^2 B}{dz^2} + \Gamma_z(z) \frac{dB}{dz} - k^2 B(z) = 0 \quad (2.49)$$

where the separation constant $k^2$ is taken to be positive, because the alternative leads to a solution that exhibits undamped oscillation as $z$ approaches infinity. (2.48) is precisely Bessel’s Equation [Abramowitz and Stegun, 1972, p. 358]:

$$\frac{1}{\xi} \frac{d}{d\xi} \left( \xi \frac{dA}{d\xi} \right) + \left( 1 - \frac{m^2}{\xi^2} \right) A(\xi) = 0 \quad (2.50)$$

with order parameter $m = 0$, and $\xi = kr$. We know that the radial electric field along the axis of an axisymmetric system must be zero. Based upon this, we demand:

$$\frac{d}{dr} A(r = 0) = 0 \quad (2.51)$$

Moreover, we have discussed in Section 2.4.2 that the correct boundary condition at the outer radial boundary is simply that the radial electric field vanishes at sufficiently large radii. The precise translation of this in the present context is

$$\lim_{a \to \infty} \left\{ \frac{d}{dr} A(a) \right\} = 0 \quad (2.52)$$

We note that condition (2.36), that the radial electric field go to zero at the simulation boundary, is an approximation to this one.

It would seem that we have identified two boundary conditions, (2.51) and (2.52) that will completely specify the solution to the second order equation (2.48). In fact, condition (2.51) rules
out the Neumann and Hankel functions as possible solutions to (2.48). This is because the behavior of the first derivative of these functions is singular at \( r = 0 \) [Abramowitz and Stegun, 1972, p. 360]. The solution is thereby narrowed to the Bessel Function of order zero:

\[
A(r) = J_0(kr)
\]

(2.53)

Now, the outer radial boundary condition would in many instances serve to specify the radial wavenumber \( k \). An example of one such instance is a condition of the form (2.36), which would specify a sequence of possible wavenumbers. The condition (2.52), however, does not determine the value of \( k \) because the first derivative of the zero order Bessel Function goes to zero in the limit of large argument, regardless of \( k \) [Abramowitz and Stegun, 1972, p. 364]. It will turn out that \( k \) must be chosen to match the low altitude and high altitude electric field solutions at the boundary between the regions in which the two solutions are applied. This boundary is defined as the \( z = z_{\text{tran}} \) plane.

In order to solve (2.49) for \( B(z) \), we rewrite the second order equation as a system of two first order, ordinary differential equations:

\[
\frac{dB}{dz} = -\Lambda \\
\frac{d\Lambda}{dz} = -k^2 - \Gamma_s(z)\Lambda
\]

(2.54a) \hspace{1cm} (2.54b)

We must identify initial values for \( B \) and \( \Lambda \) in order to integrate (2.54). The first issue to resolve is whether to specify \( (B, \Lambda) \) at the upper boundary \( (z = z_{\text{max}}) \) or the lower boundary \( (z = z_{\text{tran}}) \) of the high altitude region. We note that \( \Gamma_s(z) \) varies much more slowly than the conductivity itself, and is roughly constant over the scale length of conductivity variation. This fact is demonstrated in Figure 2.8b, using the quiescent conductivity to compute \( \Gamma \). If we consider the solution over an interval for which \( \Gamma_s(z) = \Gamma \) is constant, then the system is autonomous:

\[
\frac{d}{dz} \begin{bmatrix} B \\ \Lambda \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -k^2 & -\Gamma \end{bmatrix} \begin{bmatrix} B \\ \Lambda \end{bmatrix}
\]

(2.55)

and we can examine its eigenvalues, \( \lambda \), and eigenvectors.

\[
\lambda_\pm = -\frac{\Gamma}{2} \pm \sqrt{\left(\frac{\Gamma}{2}\right)^2 + k^2} \quad \text{with eigenvectors:} \begin{bmatrix} B \\ \Lambda \end{bmatrix} = \begin{bmatrix} 1 \\ -\lambda_\pm \end{bmatrix}
\]

(2.56)

From an examination of (2.55) and (2.56) it is apparent that the \((B, \Lambda)\) phase plane is characterized by a single stationary point at the origin, which is a saddle point. Representative
Figure 2.8: The high altitude electric field solution. (a) A composite plot of the three solutions (in regions labeled I, II, and III) that determine the electric field used in the simulation. Shown is the vertical component of the electric field along the system axis. Broken curve indicates portion of electric field profile determined from (2.54). Lower horizontal dashed line is at altitude \( z_{\text{trans}} \), upper is at \( z_{\text{max}} \). (b) Altitude gradient of logarithm of conductivity, \( \Gamma_z \), defined by (2.42). Demonstrates that \( \Gamma_z \) is approximately stepwise constant. Larger values in the 70 km to 80 km altitude range correspond to layer of rapidly increasing electron density. (c) Representative trajectories in the (B, A) phase plane, obtained from (2.54) in the case of constant \( \Gamma_z = 0.2, k = 0.2 \). System moves along trajectory toward asterisks with increasing \( z \). Dashed trajectory lies along the separatrix, leading into saddle point at position of circle.
trajectories in the \((B, \Lambda)\) phase plane are plotted in Figure 2.8c. The origin corresponds to the zero field, zero potential state at \(z \rightarrow +\infty\). Thus, if we began the integration at \(z = z_{\text{trans}}\) and proceeded upward toward \(z = z_{\text{max}}\), we would approach a saddle point as illustrated by the lower trajectory in Figure 2.8c. As the solution missed the saddle point and marched on, it would diverge toward unphysically large values of the electric field. The physically accurate solution lies along the separatrix, but an infinitesimal deviation in the initial conditions applied at \(z = z_{\text{trans}}\) will lead the system to diverge from this trajectory. We therefore prefer to integrate (2.54) beginning at the high altitude terminus—near the saddle point—and proceed downward. In this manner we will obtain a reliably convergent solution.

Figure 2.8a illustrates the three altitude regions relevant to the electric field solution. Region I, below the transition altitude, \(z_{\text{trans}}\), contains the electrostatic potential \(\Phi(z)\) solved directly from Poisson’s Equation as discussed above. Region III, above the upper limit of the simulation volume, \(z_{\text{max}}\), includes the point at infinity. Finally, in Region II we apply the quasineutral solution obtained by integrating (2.54). An important point is that the solution in Region II never arrives at the zero field stationary point. For this reason we cannot use the zero field state as an initial condition for the integration of (2.54). Instead we must choose an initial condition away from the stationary point, but nevertheless is consistent with the Region III solution. That solution has a unique trajectory in the \((B, \Lambda)\) phase plane, namely the one along the separatrix shown in Figure 2.8c, terminating at the saddle point.

In developing the electric field solution in Region III, we note that at the highest altitudes in the simulation volume, the conductivity is approaching a large, fixed value. We therefore assume that:

\[
\Gamma_z(z) = 0 \quad \text{for} \quad z > z_{\text{max}}, \quad \text{i.e., in Region III} \quad (2.57)
\]

This treats \(z_{\text{max}}\) as the lower boundary of a region of large, fixed conductivity. Judging by Figure 2.8b, which shows \(\Gamma_z\) small and roughly constant approaching \(z_{\text{max}}\) from below, this is a reasonable approximation. Under the condition (2.57), we readily solve (2.55) to obtain:

\[
B(z) = B(z_{\text{max}}) \exp[-k(z - z_{\text{max}})] \quad \text{for} \quad z > z_{\text{max}} \quad (2.58a)
\]

\[
\Lambda(z) = k \cdot B(z_{\text{max}}) \exp[-k(z - z_{\text{max}})] \quad \text{for} \quad z > z_{\text{max}} \quad (2.58b)
\]

The Region III solution in Figure 2.8a is of precisely this form, with \(k = 0.09\). We conclude from (2.58) that the initial values to be used in the integration of (2.54) are:
\[ B(z_{\text{max}}) \text{ and } \Lambda(z_{\text{max}}) = k B(z_{\text{max}}) \]  

(2.59)

where \( k \) and \( B(z_{\text{max}}) \) are actually free parameters. It remains for us to explain the manner in which we determine \( k \) and \( B(z_{\text{max}}) \).

The physical constraints on the solution in Region II are the vertical electric field \( E_z(0, z_{\text{trans}}) \) at the transition altitude and the electrostatic potential drop across the region, given by:

\[ \Delta \Phi = V_0 - \Phi(0, z_{\text{trans}}) \]  

(2.60)

where \( V_0 \) is the fixed ionospheric potential from (2.34), and \( \Phi(r, z) \) is the Poisson solution throughout the simulation volume. The vertical electric field at the transition altitude is likewise taken from the Poisson solution. We have defined the system (2.54) in such a way that \( B(z) \), and therefore \( \varphi(r, z) \) is zero as \( z \) approaches infinity. Thus, our conditions on \( \varphi(r, z) \) at the transition altitude are:

\[ \varphi(0, z_{\text{trans}}) = A(0) B(z_{\text{trans}}) = -\Delta \Phi \]  

(2.61a)

\[ -\frac{\partial}{\partial z} \varphi(0, z_{\text{trans}}) = -A(0) \frac{d}{dz} B(z_{\text{trans}}) = E_z(0, z_{\text{trans}}) \]  

(2.61b)

Substituting for \( A(r) \) and \( B(z) \) using (2.53) and (2.54), and using \( J_0(0) = 1 \) [Abramowitz and Stegun, 1972, p. 360], we rewrite (2.61) as:

\[ B(z_{\text{trans}}) = -\Delta \Phi \text{ and } \Lambda(z_{\text{trans}}) = E_z(0, z_{\text{trans}}) \]  

(2.62)

We have thus specified the solution to (2.54) in terms of known quantities at the transition altitude. However, we have found that the preferred means of integrating (2.54) is by starting at \( z_{\text{max}} \) and proceeding downward toward \( z_{\text{trans}} \). Moreover, we have specified the solution at the upper altitude limit in terms of two free parameters (2.59). We therefore will use (2.59) as initial conditions, integrate downward, and systematically vary \( k \) and \( B(z_{\text{max}}) \) to match the low altitude conditions (2.62).

The resulting solution

\[ \varphi(r, z) = J_0(kr) B(z) \]  

(2.63)

while precisely matching the Poisson solution, \( \Phi(r, z) \), in electrostatic potential and vertical and radial electric field at the position \( r = 0, z = z_{\text{trans}} \), may not match at all radii. It turns out that any vertical discontinuity in the electric field at the transition altitude is extremely small and primarily
affects the radial component. It therefore does not lead to perceptible charge accumulation at the transition plane.

2.5 Collisional Effects Model

Here we examine the sources and sinks of particles and energy that affect the electrons and ions in the multiple-fluid model. Such sources and sinks can be the result of collisional, chemical, and environmental effects. By collisional effects, we refer to the results of electron collisions with the ambient neutrals (N\textsubscript{2}, O\textsubscript{2}, and O) described in Section 2.2.1. These collisions can lead to impact ionization of the neutrals, producing additional ions and free electrons. They can also lead to dissociative attachment of electrons to molecular oxygen [Schulz, 1962]:

\[ O_2 + e^- \rightarrow O + O^- \]  

(2.64)

This process leads to the loss of electrons and the production of negative atomic oxygen ions. Finally, both collisional excitation and ionization lead to the exchange of energy between the electron and neutral populations.

2.5.1 Collision Rates

We are overwhelmingly concerned with the effects of electron collisions with the neutral constituents of the mesosphere. Collisions of electrons with other charged particles in this weakly ionized medium are so infrequent as to be negligible. At the representative altitude of 70 km, we see from Figures 2.2 and 2.3 that the ratio of charged particle to neutral particle concentrations is roughly $10^{-12}$. Even at an altitude of 100 km, the ratio remains infinitesimal, $10^{-8}$. Thus, any given electron is vastly more likely to interact with a neutral atom or molecule than with another electron.

Appendix A catalogs the many collisional processes that are incorporated in our model. Each process is characterized as either elastic, in which the total kinetic energy of the colliding particles is conserved, or inelastic, in which the electron loses energy to various excitations of the neutral particle. It is evident that there is a great variety of inelastic collisional processes to be accounted for, including rotational, vibrational, and electronic excitation, as well as outright ionization of the neutral. Energy dependent collisional cross sections were obtained for each of the processes listed in Appendix A [Wadzinski and Jasperse, 1982].
Given the collisional cross section associated with process $i$ as a function of electron velocity, $q_i(\vec{v})$, and given the velocity distribution of electrons $f(\vec{v})$, we proceed to calculate the total number of collisions of type $i$ experienced by all electrons in a unit volume, per unit time as [Gurevich, 1978, p. 77]:

$$ R_i(\vec{r}) N_i(\vec{r}) n_e(\vec{r}) = \int N_i(\vec{r}) q_i(\vec{v}) \nu n_e(\vec{r}) f(\vec{v}) \, d^3\vec{v} $$

(2.66)

where $R_i$ is a rate coefficient, and $N_i$ is the density of the neutral particles involved in process $i$. The integration indicated in (2.66) is carried out over all of velocity space, at a particular instant in time. The associated power loss by the electron population, per unit volume is:

$$ P_{\text{loss}}(\vec{r}) = \sum_i \varepsilon_i R_i(\vec{r}) N_i(\vec{r}) n_e(\vec{r}) $$

(2.67)

where $\varepsilon_i$ is the energy associated with process $i$.

We note that the local approximation is implicit in (2.66). That is, the parameters of the electron distribution function, $f(\vec{r}, \vec{v})$, vary slowly enough in the spatial coordinate, $\vec{r}$, that we may write:

$$ \bar{n}_e f(\vec{r}, \vec{v}) = n_e(\vec{r}) f(\vec{v}) $$

(2.68)

where $\bar{n}_e$ is a system average electron density. In our case, slowly enough is gauged relative to the linear dimension of one spatial cell in the simulation volume, roughly 1 km. The distribution function is normalized such that:

$$ \bar{n}_e f(\vec{r}, \vec{v}) d^3\vec{r} d^3\vec{v} = \text{number of electrons in the volume } d^3\vec{r} d^3\vec{v} $$

(2.69)

or, equivalently,

$$ \int f(\vec{v}) d^3\vec{v} = 1 $$

(2.70)

Moreover, we should point out that while time is not identified explicitly as an independent variable in these equations, all of the rates, densities, and distributions do vary in time. The time variable is omitted for simplicity of notation and for the reason that the equations will be evaluated for density and distribution functions given at particular instants in time.

### 2.5.2 Electron Distribution Function

Of course, we do not actually know the electron distribution function appearing in (2.66). It is the premise of the fluid approach that we can adequately characterize the medium by moments of the distribution function. By the same token, the fluid solution does not completely specify the
distribution function, only its moments. We are therefore left to postulate the precise form of the electron distribution function in order to carry out the collision integral in (2.66). We will examine two approaches to specifying the distribution function. First, we will propose that the electron distribution can be adequately characterized in the local approximation by a Maxwellian velocity distribution of appropriate temperature, \(T_e\) [Krall and Trivelpiece, 1986, p. 5].

\[
\bar{n}_e f(\vec{r}, \vec{v}) = n_e(\vec{r}) f_M(\vec{v}) \quad \text{with} \quad f_M(\vec{v}) = \left( \frac{m}{2\pi \kappa T_e} \right)^{3/2} \exp\left( -\frac{mv^2}{2\kappa T_e} \right) \quad (2.71)
\]

where \(m\) is the electron mass, and \(v\) is the electron speed. Alternatively, we can make use of a velocity distribution of the Druyvestein type [Gurevich, 1978, p. 70], for the known, local value of the electric field, \(E\):

\[
f_D(\vec{v}) = C \exp\left( -\int_{\kappa T + 2e^2 E^2 / 3mv_m (\delta v_m + R_s)} \right) \quad (2.72)
\]

where \(C\) is a normalization constant, \(T\) is neutral temperature, \(v_m\) is the frequency of electron-neutral momentum-transfer collisions, as a function of electron speed, \(\delta\) is the fraction of electron energy lost per elastic collision, and \(R_s\) is the fractional power lost to inelastic excitation of the ambient neutrals, also as a function of electron speed. In the absence of an electric field, and under equilibrium conditions, in which \(T = T_e\), (2.72) reduces to a Maxwellian. Both distributions are isotropic. We will consider each, in turn.

The Maxwellian distribution accurately characterizes an ideal gas in thermodynamic equilibrium with its container, under the influence of binary collisions. While the overall simulation system is surely not in a long term state of thermodynamic equilibrium, given the effect of the electric field on the electron population, it may nevertheless be reasonable to approximate small parcels of the electron population as being in a continuously, but slowly, varying state of equilibrium with their neutral gas environment. This assertion rests on the fact that we are considering a weakly ionized medium, in which the electron-neutral collision period is several orders of magnitude shorter than the millisecond time scale of macroscopic change in the system (Table 2.1). Interactions between electrons and neutral particles are mediated by short range forces, so that the duration of each collision is much shorter than the time between collisions [Krall and Trivelpiece, 1986, p. 312]. This is the definition of a binary collision. Thus, the electron population is in continuous, intimate contact with the neutral gas, under the influence of binary collisions. To first order, the neutrals are unaffected by the electric field.
We are thereby motivated to suggest that the Maxwellian velocity distribution (2.71) may satisfactorily approximate the actual electron velocity distribution. The effect of the electric field on the distribution is taken into account by varying the electron temperature, which parameterizes (2.71). The electron temperature is determined self-consistently, under the influence of the electric field, using the electron energy transport equation (Section 2.7.3). In principle, (2.71) could also be modified to account for the electron drift caused by the electric field. However, even for large electric fields on the order of 1 V/m, the drift speed is so much smaller than the electron thermal speed that the distribution function is essentially unchanged by the presence of the drift. For example, at the representative altitude of 70 km and electron temperature of 0.2 eV, the electron mobility is \( \sim 6.0 \times 10^3 \, \text{m}^2/\text{V} \cdot \text{s} \) [Kraall and Trivelpiece, 1986, p.321]. For an electric field of 1 V/m, this translates into an electron drift speed of 6,000 m/s, versus a thermal speed of 270,000 m/s. As a result, we leave the Maxwellian in the isotropic form (2.71) in spite of the anticipated drifts.

For the isotropic electron velocity distribution functions under consideration, the collision rate integral (2.66) can be written:

\[
R_i(\bar{\tau}) = 4\pi \int_0^\infty v^3 q_i(v) f(v) \, dv
\]  

(2.73)

Writing both (2.71) and (2.73) in terms of the normalized electron kinetic energy, \( \bar{\varepsilon} \):

\[
\bar{\varepsilon} = \frac{mv^2}{2\kappa T_e}
\]  

(2.74)

we obtain

\[
f_M(\bar{\varepsilon}) = \left( \frac{m}{2\pi \kappa T_e} \right)^{\frac{3}{2}} \varepsilon^{\frac{1}{2}} e^{-\varepsilon}
\]  

(2.75)

and

\[
R_i(\bar{\tau}) = 8\pi \left( \frac{\kappa T_e}{m} \right)^2 \int_0^\infty \bar{\varepsilon} q_i(\bar{\varepsilon}) f(\bar{\varepsilon}) \, d\bar{\varepsilon}
\]  

(2.76)

where \( q_i(\bar{\varepsilon}) = q_i(\kappa T_e \bar{\varepsilon}) \), and we now consider the cross section to be a function of electron energy. Combining (2.75) and (2.76), we obtain the result actually used to compute the collision rate coefficients under the assumption of a Maxwellian electron velocity distribution:

\[
R_M(\bar{\tau}) = 2\sqrt{\frac{2}{\pi}} \left( \frac{\kappa T_e}{m} \right)^{\frac{1}{2}} \int_0^\infty \bar{\varepsilon} q_i(\bar{\varepsilon}) e^{-\bar{\varepsilon}} \, d\bar{\varepsilon}
\]  

(2.77)

This would be the end of the story regarding collision rates if the cross sections of interest all peaked at energies below the electron thermal energy. In that case, the sole contribution to the
integral (2.77) would arise from the central bulk of the Maxwellian, and only if the Maxwellian assumption were grossly inaccurate would that assumption lead to a significant error in the result of the integration. However, the excitation and ionization cross sections listed in Table A.1 exhibit peaks at energies as high as tens of electron volts. These peaks are located on the high energy tail of the electron velocity distribution, giving the tail a disproportionately great effect on the calculation of the collision rates. As a result, the excitation and ionization rates depend sensitively on the exact form of the tail of the distribution, and we are compelled to give close attention to any departure from Maxwellian due to the electric field.

As discussed above, in order to compute an electron distribution function that accounts for the presence of an electric field, we must depart from a fluid approach. Instead, we must adopt a kinetic theory approach in which the electron population is characterized statistically by a velocity and spatial distribution function. In the case of a weakly ionized medium the electron distribution is governed by the Boltzmann Equation [Krall and Trivelpiece, 1986, p. 311]:

\[ \frac{\partial}{\partial t} f + \vec{v} \cdot \vec{\nabla} f - \frac{e}{m} (\vec{E} + \vec{v} \times \vec{B}) \cdot \frac{\partial}{\partial \vec{v}} f = \left. \frac{\partial f}{\partial t} \right|_c \]  

(2.78)

where the third term accounts for the acceleration experienced by an under the influence of both electric and magnetic fields, and the right hand side represents the rate of change of the distribution function due to collisions—the Boltzmann collision integral.

Gurevich [1978] demonstrates that under the circumstances of interest to us, the Boltzmann Equation can be solved analytically to obtain the Druyvestein distribution (2.72). We will outline the derivation in order to establish the assumptions that are made. In the first place, the actual solution to (2.78) is expanded in a manner that separates the isotropic and directional parts:

\[ f(\vec{v}, \vec{r}, t) = f_o(\vec{v}, \vec{r}, t) + \frac{\vec{v}}{v} \cdot \vec{f}_i(\vec{v}, \vec{r}, t) \]  

(2.79)

This expansion assumes that the departure of the velocity distribution from isotropic is small. It is particularly suited to the case in which the principal perturbation is a drift, and the drift speed is much smaller than the thermal spread of the distribution. As discussed above, the present case satisfies this assumption quite accurately.

Substituting (2.79) in (2.78), we obtain four equations for the four unknowns, namely \( f_0 \) and the three components of \( f_i \) [Gurevich, 1978, p. 58].
\[
\frac{\partial f_0}{\partial t} + \frac{v}{3} \nabla \cdot \vec{J}_1 - \frac{e}{3mv^2} \frac{\partial}{\partial v} \left( v^2 \vec{E} \cdot \vec{J}_1 \right) = -S_0 \quad (2.80a)
\]

\[
\frac{\partial \vec{J}_1}{\partial t} + v \nabla f_0 - \frac{e}{m} \frac{\partial f_0}{\partial v} - \frac{e}{m} \vec{B} \times \vec{J}_1 = -\vec{S}_1 = -\nu_m(v) \vec{J}_1 \quad (2.80b)
\]

where the form of the first order Boltzmann collision integral on the far right of (2.80b) assumes that electron-electron and electron-ion collisions may be ignored. In the weakly ionized plasma under consideration, this assumption is justified. Furthermore, under the assumption that the zero order distribution function is uniform in space, we find that (2.80b) is solved by [Gurevich, 1978, p. 65]:

\[
\vec{J}_1 = -\vec{u}(v, t) \frac{d}{dv} f_0(v) \quad (2.81)
\]

provided that \( \vec{u}(v, t) \) satisfies:

\[
\frac{\partial \vec{u}}{\partial t} = -\frac{e}{m} \left( \vec{E} + \vec{u} \times \vec{B} \right) - \nu_m(v) \vec{u} \quad (2.82)
\]

which in the steady state describes the mean electric field drift of the electrons. The assumption of uniform \( f_0 \) is valid in the local approximation.

Inserting (2.81) in (2.80a) provides:

\[
\frac{\partial f_0}{\partial t} + \frac{e}{3mv^2} \frac{\partial}{\partial v} \left[ v^2 \vec{E} \cdot \vec{u} \frac{\partial f_0}{\partial v} \right] + S_0 = 0 \quad (2.83)
\]

We evaluate the zero order collision integral by once again recalling that in a weakly ionized plasma, only electron-neutral collisions are significant. Under this assumption, we rewrite (2.83) as:

\[
\frac{\partial f_0}{\partial t} + \frac{1}{v^2} \frac{\partial}{\partial v} \left[ \frac{e}{3m} \vec{E} \cdot \vec{u} \frac{\partial f_0}{\partial v} - \frac{1}{2} \left[ \delta v_m(v) + R_e(v) \right] \left[ \frac{\kappa T}{m} \frac{\partial f_0}{\partial v} + vf_0 \right] \right] = 0 \quad (2.84)
\]

We solve (2.82) for

\[
\vec{u} = -\frac{e\vec{E}}{m\nu_m(v)} \quad (2.85)
\]

in the steady state and in the absence of a magnetic field. Using (2.85) in (2.84) and again assuming steady state conditions allows us to integrate (2.84) directly to obtain the Drayvestein distribution (2.72). The assumption of steady state conditions is valid because we are concerned with the evolution of the system over intervals much longer than the electron-neutral collision period.
In order to evaluate \(2.72\), we must specify the effective differential cross section for elastic, electron-neutral collisions in the lower ionosphere in analytical form [Gurevich, 1978, p. 62] and thereby obtain a momentum transfer collision frequency of:

\[
\nu_m = C^* N \nu^{7/5}
\]  
(2.86)

where \(C^* = 1.62E-23\) in MKS units and \(N\) is total neutral density. In addition, we need the fractional rate of energy loss by an electron due to inelastic collisions, \(R_e\). At low electron energies (< 0.5 eV) the principal inelastic loss mechanism is the rotational excitation of neutral nitrogen and oxygen [Gurevich, 1978, p. 70]. An analytic expression may be obtained in the case that the energy loss per collision is much less than the average kinetic energy of the electrons:

\[
R_e(\nu) = \frac{8B_0 g_\nu N}{m\nu}
\]  
(2.87)

where \(B_0\) is the energy of a rotational transition, 0.2E-03 eV for both molecular oxygen and nitrogen [Wadzinski and Jasperse, 1982, p. 50], and \(q_\nu\) is the associated cross section, which is roughly constant over energy at 2.0E-20 m². Indeed, we see that

\[
B_0 << \kappa T
\]  
(2.88)

for a neutral temperature of roughly 25 meV, justifying the assumption that went into \(2.87\).

Finally, the fractional energy lost in a momentum transfer collision is just [Wadzinski and Jasperse, 1982, p. 46]:

\[
\delta = \frac{2m}{M}
\]  
(2.89)

where \(M\) is the mass of the neutral particle. For molecular nitrogen or oxygen, this takes a value of approximately 3.5E-05.

To preliminarily assess the relative importance of the electric field-induced departure from a Maxwellian, we compare the magnitudes of the two terms in the denominator of the exponent of \(2.72\), recalling that when the second term is negligible, the distribution reduces to a Maxwellian. We use representative numbers for an altitude of 70 km (chosen because at this altitude we observe both significant heating and significant electron densities), namely neutral density and electron-neutral momentum transfer collision frequency of \(10^{21}\) per m³ and \(10^7\) Hz, respectively. For an electric field of 1 V/m and an electron temperature of 0.2 eV we find:

\[
\kappa T << \frac{2e^2E^2}{3m\nu(\nu)} \approx 0.6\text{ eV}
\]  
(2.90)
In this limit, the electric field has an important effect on the velocity distribution. Moreover, we may ignore the neutral temperature term and integrate (2.72) analytically to obtain a definite expression for the Druyvestein velocity distribution, which we then normalize according to (2.70), yielding:

\[
f_D(v) = \frac{2}{3\pi \alpha^3 \Gamma(\alpha)} \exp \left[ -\left( \frac{v}{\alpha} \right)^{\frac{2}{3}} \right], \text{ where } \alpha^{\frac{1}{3}} = \frac{2e^2 E^2}{9mC^*N^2B_0q}, \tag{2.91}
\]

where \( \Gamma(x) \) is the gamma function [Abramowitz and Stegun, 1972, p. 255].

We point out that (2.91), unlike the Maxwellian, is parameterized not by the electron temperature, but by the quasi-static electric field strength. In fact, we can calculate an effective temperature, \( T_{ef} \), from (2.91) using the average kinetic energy of the electrons for an isotropic velocity distribution [Krall and Trivelpiece, 1986, p. 84]:

\[
\kappa T_{ef} = \frac{2}{3} \int f(\vec{v}) \left( \frac{mv^2}{2} \right) d^3\vec{v} \tag{2.92}
\]

For the case of a Maxwellian velocity distribution, \( T_{ef} = T \). Using (2.91) in (2.92), we find

\[
\kappa T_{ef} = \frac{ma^2}{3} = \frac{m}{3} \left( \frac{2e^2 E^2}{9mC^*N^2B_0q} \right)^{\frac{3}{4}} \tag{2.93}
\]

so that the effective temperature of the electrons is itself a function of the electric field. The practical impact of employing the kinetic theory approach is that the Boltzmann Equation has incorporated the effects of the electric field and elastic electron-neutral collisions on the zero order electron velocity distribution. To a less adequate degree, the effects of inelastic collisions have been incorporated as well. In this manner, collisional and electric field effects on electron temperature have been built into the Druyvestein distribution (2.91), and this approach has circumvented the electron energy transport equation that determines the electron temperature in the fluid theory.

We can now use the effective temperature in place of the parameter \( \alpha \) in the Druyvestein distribution:

\[
f_D(v) = \frac{2}{3\pi \Gamma(\alpha)} \left( \frac{m}{3\kappa T_{ef}} \right)^{\frac{3}{2}} \exp \left[ -\left( \frac{mv^2}{3\kappa T_{ef}} \right)^{\frac{2}{3}} \right] \tag{2.94}
\]

which we may rewrite in terms of the normalized electron kinetic energy, \( \bar{\epsilon} \), defined in (2.74):
\[ f_D(\varepsilon) = \frac{2}{9\pi^{3/2}} \left( \frac{m}{\kappa T_e} \right)^{3/2} \Gamma(\kappa) \exp \left\{ -\left(\frac{2}{3} \varepsilon \right)^{3/2} \right\} \]  

(2.95)

For purposes of comparison, we reproduce the corresponding form of the Maxwellian distribution (2.75):

\[ f_M(\varepsilon) = \frac{1}{2\pi^{3/2}} \left( \frac{m}{\kappa T_e} \right)^{3/2} \exp \left\{ -\varepsilon \right\} \]  

(2.96)

We compare the shapes of the two distributions in Figures 2.9d and 2.9e. From the linear plot of Figure 2.9d, it is apparent that the absolute magnitudes of Druyvestein and Maxwellian distributions are quite similar at all energies, but the logarithmic plot of Figure 2.9e demonstrates that they nevertheless differ by a large and growing factor at high energies. Their ratio is roughly unity at energies on the order of the thermal energy, but orders of magnitude in the high energy tail, where the absolute magnitude of both distributions is extremely small. The Druyvestein falls off faster with increasing energy than does the Maxwellian. This is a result of the fact that the loss process considered in obtaining the Druyvestein, namely rotational excitation of ambient neutrals, has a cross section that is flat over the energy range of interest, so that energy is removed preferentially from the faster, high energy electrons due to their more frequent collisions.

The ultimate effect of the discrepancy between the high energy behaviors of the two distributions is simply this. Collisional processes with cross sections that peak at energies much greater than the thermal energy of the electrons will proceed more slowly for electrons obeying the Druyvestein than for those obeying the Maxwellian. Increasing temperature reduces the difference between the two cases, because at higher temperatures the energy axis is rescaled in such a way that both distributions expand toward higher energies. Viewed another way, a cross section peaking at some fixed energy will find itself closer to the central core of either distribution as temperature increases. In this region of the distribution, the discrepancy loses significance.

2.5.3 Collisional Effects Calculated For Maxwellian Versus Druyvestein Electron Distributions

Decreasing discrepancy at higher temperatures is precisely the trend we see at work in Figures 2.9a through 2.9e, which show the rates of collisional effects computed for the Druyvestein velocity distribution compared to those for the Maxwellian. Figure 2.9a shows the sum of the
Figure 2.9: Collisional effects calculated for Maxwellian (M) versus Druyvestein (D) electron velocity distributions. Neutral environment taken for 70 km altitude. Abscissa in plots (a) - (c) is electron temperature for Maxwellian case and effective temperature for Druyvestein case. (a) Frequency of electron-neutral, momentum transfer collisions experienced by one electron, showing no difference between results for the two distributions. (b) Energy loss rate for one electron due to all collisional processes. Bump in loss curve between 0.2 and 2.0 eV is due to vibrational excitation. Horizontal broken line points out the small magnitude of the difference in temperature associated with a given loss rate. (c) Net ionization rate per electron per second, showing the significant discrepancy between results for the two distributions at temperatures below 6 eV. (d) and (e) Comparison of Maxwellian and Druyvestein energy distributions, showing that they differ by a small factor at low energies, but a huge factor at high energies. Increasing temperature rescales the energy axis in such a way that any absolute energy moves toward the center of the distributions.
momentum transfer collision rates with the three neutral species at 70 km altitude. For the
Maxwellian case, this is computed using \( R_{\text{M}} \) from (2.77):

\[
\nu_{\text{mom}}(\vec{r}, T_e) = \sum_{i, \text{momentum}} R_{\text{M}, i} \, N_i
\]  

(2.97)

The corresponding result for the Druyvestein distribution uses \( R_{\text{D}} \), obtained by inserting (2.95) in
(2.76):

\[
R_{\text{D}}(\vec{r}, T_e) = \frac{16}{9\sqrt{3}} \frac{\kappa T_e}{\Gamma(\%)} \left( \frac{T_e}{m} \right)^{1/2} \int_{0}^{\infty} \tilde{q}_i(\tilde{e}) \, \exp\left\{-\left(\frac{\tilde{e}}{T_e}\right)^{1/2}\right\} \, d\tilde{e}
\]  

(2.98)

The two curves of momentum transfer collision frequency are virtually indistinguishable because
the momentum transfer cross sections are not strongly peaked in energy, so that differing
contributions from the high energy tails are swamped by large, similar contributions from the
central core of the distributions.

Figure 2.9b shows the mean rate of energy loss by an electron due to all collisional
processes that are modeled, both elastic and inelastic. Again we compare the results obtained using
the two alternative velocity distributions. The energy loss rate per electron is computed from the
rate coefficients \( R_{\text{M}} \) and \( R_{\text{D}} \) using:

\[
\Phi_{\text{M,D}}(\vec{r}, T_e) = \sum_{i, \text{all}} \Delta \epsilon_i \, R_{\text{M,D}, i} \, N_i
\]

(2.99)

where \( \Delta \epsilon_i \) is the energy loss associated with each excitation, ionization, or elastic collisional
process. We observe that there is a narrow range of electron temperatures, 0.2 to 0.3 eV, for
which the loss rate calculated for the Maxwellian appreciably exceeds that for the Druyvestein.
The discrepancy arises at these temperatures because the central core of each velocity distribution
is positioned such that the overlap of the high energy tails with the sharp onsets of the vibrational
cross sections accentuates the difference between the distributions. At lower temperatures, elastic
losses dominate. At higher temperatures, the large vibrational cross sections move toward the core
of the distributions and overwhelm the effects of the ionization cross sections.

Some degree of horizontal discrepancy between the curves in Figure 2.9b exists over the
range 1.0E+04 to 1.0E+06 eV per s in energy loss rate. However, the maximum difference
between electron temperatures on the two curves at a fixed level of energy loss is less than 20%. In
other words, the use of a Maxwellian will cause the temperature associated with a known energy
loss rate to be underestimated by less than 20%. This degree of inaccuracy cannot be considered
significant, particularly in view of Table A.1, which shows that the range of electron temperatures
affected, namely 0.2 to 0.3 eV, is not in the vicinity of major excitation or ionization thresholds. We are assessing the significance of the observed discrepancy in an appropriate manner, based upon the fact that the energy loss rate will be used to determine electron temperature from an energy balance equation (see Section 2.7.3).

Figure 2.9c shows the sum of the rates of all the ionization processes, using neutral densities that are appropriate for 70 km altitude. The rate of each ionization process is calculated in a manner similar to (2.76), except that the integrand must be weighted by a factor to account for the fact that not all ionizing collisions produce an ion-electron pair. For impacts by electrons whose energy is near the ionization potential of the neutral particle, both the primary and the secondary electrons have very little energy subsequent to the ionization, and one of them will be recaptured by the newly created ion within nanoseconds. To account for this effect, the net ionization rate coefficient associated with process $i$ is written as [Strickland et al., 1983]:

$$g_i(F, T_e) = 8\pi n_\epsilon(F)\left(\frac{\kappa T_e}{m}\right)^{1/2} \int_0^{\infty} \frac{\kappa T_e}{e \varphi_{wi}} \tilde{e} - 1 \tilde{e} q_i(\tilde{e}) f(\tilde{e}) d\tilde{e}$$

(2.100)

where $\varphi_{wi}$ is the work function associated with process $i$. Since the ionization cross section is zero for electron energies below $e \varphi_{wi}$, the integral (2.100) never evaluates to a negative number. The total ionization rate is then:

$$G_{M,D}(F, T_e) = \sum_{i, \text{ion}} g_{i,M,D} N_i$$

(2.101)

Because the ionization cross sections have sharp thresholds in energy, we see from Figure 2.9c that the differing high energy tails of the two distributions result in significantly different ionization predictions when the electron temperature is much below the ionization energy. A calculation using the Maxwellian distribution will overestimate the ionization rate at low temperatures. To be specific, the ratio of the ionization rate predicted using the Maxwellian to that predicted using the Druyvesteine is huge at temperatures below several eV, falls to 2.0 at an electron temperature of 6.3 eV and drops rapidly at higher temperatures. We will regard 6.5 eV as the maximum temperature for which the discrepancy is significant. Moreover, at temperatures somewhat below this, only insignificant rates of ionization are predicted even using the Maxwellian. There is thus a window of electron temperature in which the difference in the velocity distributions is important to the simulation results.

We quantify the lower limit of this temperature window by examining the rate of ambient background ionization. If the predicted ionization rate is only a small fraction ($< 10\%$) of the
background rate, then it is not significant. The ambient ionization at nighttime is due to galactic cosmic rays. The cosmic ray ionization rate, normalized to standard temperature and pressure (273°K and 1000 mb, respectively) is measured to be 3.0E+08 per m³·s [Rosen and Hofmann, 1981b]. At the neutral pressure and temperature characteristic of 70 km, i.e., 5.72E-02 mb and 209.4°K, respectively [COSPAR, 1972], the corresponding ionization rate is 2.24E+04 per m³·s. At a typical nighttime electron density of 1.0E+08 per m³ (see Figure 2.11), this equates to an ionization rate of 2.2E-04 per second per electron. From Figure 2.9c, the ionization rate predicted for Maxwellian electrons at a temperature of 0.5 eV is 8.0E-07 per m³·s, or a tiny fraction of the background rate. At temperatures below 0.5 eV, the predicted ionization rate is not significant compared to the background rate.

We have thus identified a range of electron temperatures for which a Maxwellian velocity distribution significantly overestimates the ionization rate and for which the predicted ionization rate is potentially important to the simulation:

\[ 0.5 \text{ eV} < \kappa T_e < 6.5 \text{ eV} \]  \hspace{1cm} (2.102)

At temperatures below this range, the predicted ionization is inconsequential, and above this range the thermal spread of both velocity distributions encompasses the ionization thresholds, so that the calculated ionization rates do not differ significantly depending upon the distribution employed.

### 2.5.4 Hybrid Fluid-Kinetic Approach

Based upon the above discussion comparing the collisional effects calculated using Maxwellian and Druyvesteijn velocity distributions, we conclude that neither the electron-neutral momentum transfer collision rate, nor the total electron energy loss rate is significantly affected by the choice of a Maxwellian. The fundamental reason for this is that while an electric field does alter the high energy tail of the electron distribution, the overwhelming effect is on the thermal spread of the central core of the distribution. When we view both distributions as parameterized by their effective temperatures, we see that they respond identically to this bulk heating effect. On the other hand, the sharp, high energy thresholds of the ionization cross sections magnify the influence of the high energy tail and produce a significant overestimate of ionization when the Maxwellian is employed.

We will therefore make use of collision frequencies and energy loss rates calculated from a Maxwellian electron velocity distribution whose temperature parameter is determined self-
consistently within the fluid formulation. This is important for the logical consistency of the fluid approach. We will see (Section 2.7.3) that in the fluid formulation we can fully account for the effects of inelastic processes on electron temperature by using an energy transport equation that incorporates all relevant loss processes in a loss rate term given by (2.99). In this manner, the system of fluid equations is closed, and the electron temperature as well as the rates of neutral particle excitation emerge self-consistently from the solution.

Likewise, a comprehensive kinetic theory formulation that solves Boltzmann's Equation in the general case and incorporates all relevant processes in the collision integral would lead to an accurate, self-consistent solution for the temperatures, densities and drifts of each species. However, a hybrid approach that makes use of kinetic-theory-derived distributions for calculations within a fluid framework necessarily suffers from a lack of internal consistency. This is because the distribution functions will incorporate their own, independent solutions for the fluid densities, drifts and temperatures, and these will be unrelated and generally unequal to the solutions of the fluid equations.

For example, (2.93) shows the electron temperature calculated as the second moment of the Druyvestein distribution. This result specifies the temperature as a function of the applied electric field. Our fluid equations will include an equivalent, but different, specification in the form of the steady state energy transport equation. The differences stem from the simplifying assumptions that went into the Druyvestein distribution (2.72), principally that the velocity distribution is isotropic and that an exhaustive list of high energy inelastic processes could be ignored in favor of an analytic model of rotational excitation loss.

The Druyvestein result is nevertheless extremely useful, particularly when it is rewritten parameterized by its effective temperature (2.94). In this form it is no longer sensitive to the nature of the loss processes included, for an additional or different loss process in (2.93) will change the computed value of the effective temperature, but not the form of (2.94). Thus, if we correctly compute the electron temperature through other means, such as the energy transport equation, the Druyvestein distribution parameterized by this temperature will correctly reflect electron heating. Of course, (2.94) still has no provision for electron drift, but we have pointed out that this shortcoming does not affect the intended use of the distribution, which is to calculate collisional energy loss.

Based upon the preceding discussion, we apply the following approach to modeling collisional effects. (a) We will use a Maxwellian velocity distribution to calculate electron-neutral
momentum transfer collision rate and energy loss rate according to (2.77), (2.97), and (2.99). In cases for which ionization is significant, we will use the Druyvestein velocity distribution to calculate the ionization rate according to (2.95), (2.100), and (2.101). The effective electron temperature used in the Druyvestein is obtained from the fluid equations using the energy loss rate calculated in step (a).

To avoid overestimating the ionization rate, we have resorted to a hybrid fluid/kinetic approach. We have circumvented concerns about self-consistency by ignoring the electric field dependence of the electron temperature suggested by (2.93) and instead considering temperature as a free parameter in the Druyvestein velocity distribution, to be determined by the fluid equations. In cases for which ionization is unimportant, this step is irrelevant.

This compromise allows us to preserve the overarching fluid formulation, which is desirable for the reasons outlined in Chapter 1. Foremost, the fluid approach is inherently non-local, as is the system we are simulating. Moreover, the fluid approach is ideally suited to account for large-scale redistribution of the charged particles, and such drifts will turn out to be crucially important.

As a final note in this section, we point out that dissociative attachment (2.64) is a process with a sharp energy threshold near 6 eV. As a result, we treat it as we did ionization and compute its rate per electron from the associated cross section, \(q_d\) [Schulz, 1962], in a manner analogous to (2.97) and (2.98):

\[
G_d(\vec{r}, T_e) = \frac{16}{9\sqrt{3} \Gamma(\frac{\alpha}{2})} n_e(\vec{r}) N_{O_2}(\vec{r}) \left( \frac{\kappa T_e}{m} \right)^{\frac{1}{2}} \int_0^\infty \tilde{\varepsilon} q_d(\tilde{\varepsilon}) \exp\left\{ - \left( \frac{\tilde{\varepsilon}}{\frac{3}{2}} \right)^{\frac{5}{2}} \right\} d\tilde{\varepsilon} \tag{2.103}
\]

where we have again employed the Druyvestein, by the same reasoning as for the ionization rate.

2.6 Chemical Effects Model

By ion chemistry we refer to the complex array of interactions among the many charge carrying species in the lower ionosphere. Charge carriers can be created or annihilated, in the sense that they can enter or leave the neutral population. They can also transform from one to another. All of these processes result in local changes in charge carrier concentration which are governed by a system of coupled rate equations.
2.6.1 Species, Reactions, and Rates

Even though only a small subset of the charged species—in some cases only a single species—may be responsible for the conductivity in a particular altitude regime, we must keep track of all the possible species everywhere. This is because charged populations that are not themselves electrically significant may nevertheless play critical roles in reactions that enhance or deplete more conductive species. For example, relatively immobile, large positive ions are involved in reactions that regulate the density of small positive ions, a dominant conductive species at low altitudes.

We may group the individual charged species into several classes based upon their polarity and mobility. Table 2.2 lists the classes and their dominant members. Three of the classes have been introduced already, in Section 2.2.2. They were singled out because their high mobility makes them overwhelmingly responsible for electrical conduction throughout the range of altitudes under consideration. Two additional classes, cluster ions and aerosols, are included here because they play important roles in the chemistry of the region, even if they do not contribute to the conductivity. The cluster ions, also known as large positive ions or proton hydrates, are hydronium ions, $\text{H}_3\text{O}^+$, about which are clustered varying numbers of water molecules. They are denoted by $\text{H}_3\text{O}^+\cdot(n\text{H}_2\text{O})$ or $\text{H}^+\cdot(n\text{H}_2\text{O})$, where the index $n$ indicates the level of clustering. Cluster ions as large as $\text{H}_5\text{O}_2^+$ are found in appreciable concentrations up to 80 km altitude [Swider et al., 1978]. Aerosols are macroscopic particles that range in diameter from nanometers to microns and stem from forest fires, dust storms, and volcanic eruptions. Small ions can suffer attachment to aerosols, effectively eliminating them as conductive particles.

<table>
<thead>
<tr>
<th>Class</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>electrons</td>
<td>$e^-$</td>
</tr>
<tr>
<td>small negative ions</td>
<td>$\text{NO}_3^-$ (primary), $\text{O}_2^-$</td>
</tr>
<tr>
<td>small positive ions</td>
<td>$\text{O}_2^+$ and $\text{NO}^-$ (primary), $\text{N}_2^-$</td>
</tr>
<tr>
<td>cluster ions</td>
<td>$\text{H}^+\cdot(n\text{H}_2\text{O})$</td>
</tr>
<tr>
<td>aerosols</td>
<td>smoke particles</td>
</tr>
</tbody>
</table>

Table 2.2: Classes of charged particles involved in ion chemistry.

The reactions among the charged species can all be classified as recombination, attachment, or detachment. Recombination can remove pairs of positive and negative ions, as well as of positive ions and electrons. Attachment and detachment describe the dual processes:
\[ X + e^- \leftrightarrow X^- \]

where \(X\) is a representative neutral species, including water vapor and aerosols. Chemical processes can also involve large ions, whose dynamics we do not track because they are so massive as to be immobile. At low altitudes, these large ions can nevertheless act as important sources and sinks of small ions, through the processes of hydration and dehydration.

In addition, several external environmental effects contribute to the populations of ionized species. These effects were mentioned in Section 2.2.2 on the ambient charged constituents. They include photoionization, cosmic ray ionization, and ionization due to terrestrial radioactivity. These processes depend only upon the densities of the neutral species, which are roughly constant, and upon other parameters external to the model. As a result, we anticipate that the environmental sources of the charged species will remain constant over the course of the simulation.

Environmental effects on the particle rate equations will be treated by including a steady state ionization source term.

In Table 2.3 we catalog all of the production processes relevant in the 0 to 100 km altitude range. Three body attachment involving two neutral molecules and one electron is important at these altitudes owing to the large ambient neutral densities. The process of hydration takes small positive ions and turns them into cluster ions. Because they will become part of the rate equations, we include the effects of environmental ionization, as well as electron impact ionization as calculated in Section 2.5.

<table>
<thead>
<tr>
<th>Class</th>
<th>Production Process</th>
<th>Rate</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>electrons</td>
<td>cosmic rays</td>
<td>(I) determined from equilibrium</td>
<td></td>
</tr>
<tr>
<td></td>
<td>net impact ionization</td>
<td>(G)</td>
<td>see Section 2.5.3</td>
</tr>
<tr>
<td></td>
<td>(X^- \rightarrow X + e^-)</td>
<td>(\lambda) (3.0\text{E-23 m}^3/\text{s})</td>
<td>\cite{Glukhov et al., 1992}</td>
</tr>
</tbody>
</table>

Table 2.3: Charged particle production processes. \(N\) = total neutral density. Rates for two-body processes are given in \(\text{m}^3/\text{s}\), three-body processes in \(\text{m}^2/\text{s}\). Continued.
<table>
<thead>
<tr>
<th>small negative ions</th>
<th>three body attachment</th>
<th>$e^- + O_2 + O_2 \rightarrow O_2^- + O_2$</th>
<th>$\alpha_3$</th>
<th>1.5E-41(300/T_e) $\cdot$ exp(-600/T_e) m$^3$/s</th>
<th>[Banks and Kockarts, 1973, p. 278]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dissociative attachment</td>
<td>$e^- + O_2 \rightarrow O + O^-$</td>
<td>$\alpha_D$</td>
<td>See Section 2.5.4</td>
<td></td>
</tr>
<tr>
<td>small positive ions</td>
<td>cosmic rays</td>
<td>$I$</td>
<td>determined from equilibrium</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dehydration of cluster ions</td>
<td>$NO^- (H_2O) + N_2 \rightarrow NO^- + H_2O + N_2$</td>
<td>$\lambda_C$</td>
<td>&lt; 1.0E-20 m$^3$/s</td>
<td>See Appendix B</td>
</tr>
<tr>
<td></td>
<td>net impact ionization</td>
<td>$G$</td>
<td>see Section 2.5.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cluster ions</td>
<td>hydration of small positive ions</td>
<td>$NO^- + H_2O + N_2 \rightarrow NO^- (H_2O) + N_2$</td>
<td>$B_C$</td>
<td>1.6E-40 m$^3$/s</td>
<td>[Rowe et al., 1974]</td>
</tr>
</tbody>
</table>

Table 2.3 continued: Charged particle production processes. $N = $ total neutral density. Rates for two-body processes are given in m$^3$/s, three-body processes in m$^3$/s.

The use of aggregate reaction rates for the processes listed in Table 2.3 requires some justification. A thoroughly detailed model of ion chemistry would account individually for the myriad reactions affecting each member of each ion class. For example, in the bottom row of Table 2.3 we refer to the hydration of small positive ions. In fact, we know that there are several small positive ions, and many degrees of clustering. To cite only a few of the possibilities, three-body reactions involving either $O_2$ or $N_2$ can result in the clustering of molecular oxygen ions, while a three-body reaction buffered by $N_2$ results in the hydration of nitric oxide ions [Banks and Kockarts, 1973, pp. 274, 276]

$$O_2^- + H_2O + N_2 \rightarrow O_2^- \cdot H_2O + N_2 \text{ with a rate } 2.5E-40 \text{ m}^3/\text{s} \quad (2.104a)$$

$$O_2^- + H_2O + O_2 \rightarrow O_2^- \cdot H_2O + O_2 \text{ with a rate } 2.3E-40 \text{ m}^3/\text{s} \quad (2.104b)$$

$$NO^- + H_2O + N_2 \rightarrow NO^- \cdot H_2O + N_2 \text{ with a rate } 1.5E-40 \text{ m}^3/\text{s} \quad (2.104c)$$

Subsequent reactions can create ions with an even greater degree of clustering.

Three facts regarding these reactions ease our task in simplifying this system. First, while either $O_2$ or $N_2$ can act as buffers, these neutral molecules are present in fixed proportions.
throughout the altitude range of interest. Second, the reaction rates are not widely different. Finally, the subsequent reactions involving the hydrated products of these reactions do not remove them from the class of cluster ions. As a result, Rowe et al. [1974] were able to characterize the hydration of small positive ions as a single process, described by the single-body reaction rate 1.0E-43(N^0) per s, where N is the total density of neutral particles. In this manner, a class of species is treated effectively in the aggregate.

The Rowe et al. [1974] result, however, is valid only over the narrow altitude range of 70-90 km, where the water vapor mixing ratio is roughly constant. We cannot apply it throughout the simulation volume, which extends all the way to the ground. In order to accurately account for hydration at all altitudes, we must undertake a significant elaboration of Rowe et al. [1974] and subsequent work by Glukhov et al. [1992]. The goal of this effort is to correctly determine the rate of the dehydration reaction included in our four species chemical model, in light of the fact that unmodeled reactants play an important role in this process. Appendix B explains the chain of reactions that compose the hydration process, and that leads to the calculated value of $\lambda_C$ appearing in Tables 2.3 and 2.4.

Other types of reactions, affecting other classes of particles, are treated similarly. In the case of three body attachment, the reaction listed in Table 2.3 has a rate at least an order of magnitude greater than an alternative reaction producing the same outcome [Rowe et al., 1974]:

$$e^- + O_2 + N_2 \rightarrow O_2^- + N_2 \text{ with a rate } 1.0E-43 \text{ m}^6/\text{s} \quad (2.105)$$

Thus, the reaction in the table dominates despite an almost 4:1 numerical advantage of the nitrogen buffer molecules. Finally, the only dissociative attachment process of importance involves oxygen. While molecular nitrogen can undergo resonant dissociation by electron impact—indeed, the peak cross section for this process is even greater than that for dissociative attachment to oxygen, 2.5E-22 versus 1.3E-22 m²—the resulting molecular nitrogen ion is metastable and quickly decays into two neutral atoms and a thermal electron [Spence and Burrow, 1979]:

$$e^- + N_2 \rightarrow (N_2^- \ast) \rightarrow N + N + e^- \quad (2.106)$$

While this process can be included as an electron energy loss mechanism, it results in no net loss of electrons or creation of ions and as such is left out of Table 2.3.

Table 2.4 lists all of the relevant loss processes for each class of charged particle. With the exception of recombination and aerosol attachment, these processes have already been addressed in the context of the corresponding production process. Recombination can involve electrons and ions, or pairs of oppositely charged ions. Depending upon the reactants, the rate of
recombination can vary widely. *Rosen and Hofmann* [1981b] measured the recombination coefficient for undifferentiated positive and negative ions as a function of altitude in the atmosphere. This result applies directly to the ion classes in the aggregate. The electron-small positive ion recombination coefficient is taken to be characterized by the rate for the dissociative recombination of the dominant positive ion species, nitric oxide [Swider et al., 1978]. The rate of dissociative recombination far exceeds that of radiative recombination [Wadzinski and Jasperse, 1982, p. 63]. Finally, the rate at which electrons recombine with cluster ions varies from 1.0E-12 to 1.0E-11 m³/s with increasing cluster size. The value presented in Table 2.4 is characteristic of cluster ions with \( n \) between 3 and 4, which dominate over the altitude range of interest [Reid, 1986, p.187].

<table>
<thead>
<tr>
<th>Class</th>
<th>Loss Process</th>
<th>Rate</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>electrons</td>
<td>three body attachment ( e^- + O_2 + O_2 \rightarrow O_2^- + O_2 )</td>
<td>( \alpha_3 ) 1.5E-41(300/T_e) \cdot \exp(-600/T_e) m^6/s</td>
<td>[Banks and Kockarts, 1973, p. 278]</td>
</tr>
<tr>
<td></td>
<td>dissociative attachment ( e^- + O_2 \rightarrow O + O^- )</td>
<td>( \alpha_D )</td>
<td>See Section 2.5.4</td>
</tr>
<tr>
<td></td>
<td>recombination with small positive ions ( e^- + NO^+ \rightarrow N + O )</td>
<td>( \alpha_E ) 1.58E-13 \cdot (1000/T_e)^{0.83} m³/s</td>
<td>[Wadzinski and Jasperse, 1982, p. 64]</td>
</tr>
<tr>
<td></td>
<td>recombination with cluster ions</td>
<td>( \alpha_{EC} ) 4.6E-12 m³/s</td>
<td>[Rowe et al., 1974]</td>
</tr>
<tr>
<td>small negative ions</td>
<td>detachment ( X^- \rightarrow X + e^- )</td>
<td>( \lambda ) 3.0E-23 m³/s</td>
<td>[Glukhov et al., 1992]</td>
</tr>
<tr>
<td></td>
<td>recombination with positive ions</td>
<td>( \alpha ) function of altitude</td>
<td>[Rosen and Hofmann, 1981b]</td>
</tr>
<tr>
<td></td>
<td>attachment to aerosols ( \beta )</td>
<td>5.1E-11 m³/s</td>
<td>[Rosen and Hofmann, 1981b] and references therein</td>
</tr>
</tbody>
</table>

Table 2.4: Charged particle loss processes. N = total neutral density. Rates for two-body processes are given in m³/s, three-body processes in m⁶/s. Continued.
<table>
<thead>
<tr>
<th>Process</th>
<th>Rate Constant</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>small positive ions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>hydration of small positive ions</strong></td>
<td>$B_c$</td>
<td>[Rowe et al., 1974]</td>
</tr>
<tr>
<td>$NO^+ + H_2O + N_2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table><p>ightarrow NO^+(H_2O) + N_2$ | 1.6E-40 m$^6$/s                |                                                |
| <strong>recombination with small negative ions</strong>  | $\alpha$                       | [Rosen and Hofmann, 1981b]                     |
| <strong>recombination with electrons</strong>             | $\alpha_e$                     | [Wadzinski and Jasperse, 1982, p. 64]          |
| $e^- + NO^+ \rightarrow N + O$              | 1.58E-13 \cdot (1000/T_a)^{0.83} m$^3$/s |                                                |
| <strong>attachment to aerosols</strong>                  | $\beta$                        | [Rosen and Hofmann, 1981b]                     |
|                                              | 5.1E-11 m$^3$/s                |                                                |
| <strong>cluster ions</strong>                            | <strong>recombination with small negative ions</strong> | [Rosen and Hofmann, 1981b]                     |
| <strong>dehydration of cluster ions</strong>             | $\lambda_c$                    | See Appendix B                                 |
| $NO^+(H_2O) + N_2 \rightarrow NO^+ + H_2O + N_2$ | &lt; 1.0E-20 m$^3$/s function of altitude |                                                |
| <strong>recombination with electrons</strong>            | $\alpha_{EC}$                  | [Rowe et al., 1974]                            |
|                                              | 4.6E-12 m$^3$/s                |                                                |</p>

Table 2.4 continued: Charged particle loss processes. $N = $ total neutral density. Rates for two-body processes are given in m$^3$/s, three-body processes in m$^6$/s.

The coefficient given in Table 2.4 for the attachment of ions to aerosols is representative of aerosol particles of 0.35 $\mu$m radius. It is obtained using the approximate relationship [Rosen and Hofmann, 1981b, and references therein]:

$$\beta(r_a) = 2.35E-10 \cdot r_a^{1.457} \text{ m}^6/\text{s}$$  \hspace{1cm} (2.107)

where $r_a$ is the radius of the aerosol particle in micrometers. Since the aerosol particles are distributed over several orders of magnitude in size [Gringel et al., 1986, p.170], the correct means of calculating the aerosol attachment rate is [Rosen and Hofmann, 1981b]:

$$\text{ion-aerosol attachment rate} = n_X \int_0^\infty \beta(r_a) \gamma(r_a) \, dr_a$$  \hspace{1cm} (2.108)

where $n_X$ is the density of the ions in question, and $\gamma(r_a)$ is the differential size distribution of the aerosol particles. Both the size distribution and the altitude profile of the atmospheric aerosol content are wildly variable, depending on such influences as forest burning in the Amazon, sand storms in the Sahara, and volcanic eruptions in the South Pacific [Gringel et al., 1986, p.168].

Since this variability will overwhelm any systematic inaccuracy caused by the selection of a
particular, representative value of \( r_n \), and since aerosol attachment is not expected to be significant above 10 km altitude in any case [Rosen and Hofmann, 1981b], we are justified in choosing the representative value \( r_n = 0.35 \mu m \) mentioned above. This value is intended to match the ion-aerosol attachment rate obtained by Rosen and Hofmann [1981b] at an altitude of 20 km. Given the aerosol particle size, the attachment rate is now computed from:

\[
\text{ion-aerosol attachment rate } = n_N N_A \beta
\]  

(2.109)

where \( N_A \) is the aerosol concentration.

### 2.6.2 Rate Equations and Equilibrium Ion Densities

Having now established the rate coefficients that quantify all of the chemical and collisional processes that affect the electron and ion densities in the system, can write a set of four coupled rate equations to govern the evolution of \( n_e, n_m, n_p, \) and \( n_z \), the electron, small negative ion, small positive ion, and cluster ion densities, respectively.

\[
\frac{d n_e}{dt} = I + G + \lambda n_m N - \alpha_3 n_e (N_{O2})^2 - \alpha_D N_{O2} n_e - \alpha_E n_p n_e - \alpha_{EC} n_c n_e - \beta N_A n_e
\]  

(2.110a)

\[
\frac{d n_p}{dt} = I + G + \lambda_c N_{N2} n_e - B_c N_{N2} N_r n_p - \alpha n_p n_m - \alpha_E n_p n_e - \beta N_A n_p
\]  

(2.110b)

\[
\frac{d n_m}{dt} = \alpha_3 n_e (N_{O2})^2 + \alpha_D N_{O2} n_e - \alpha n_m (n_p + n_e) - \lambda n_m N - \beta N_A n_m
\]  

(2.110c)

\[
\frac{d n_z}{dt} = B_c N_{N2} N_r n_p - \lambda_c N_{N2} n_c - \alpha n_m n_c - \alpha_{EC} n_c n_e - \beta N_A n_c
\]  

(2.110d)

where \( N \) is the total neutral density, and \( r_w \) is the water vapor mixing ratio. While the coefficients in (2.110) involve either fixed reaction rates or locally fixed neutral densities, they are generally functions of altitude. In fact, the several neutral densities, the aerosol density, the water vapor mixing ratio, and the ion-ion recombination rate should all be written:

\( N(z), N_{O2}(z), N_{N2}(z), N_A(z), r_w(z), \) and \( \alpha(z) \), respectively. The altitude dependence is omitted from (2.110) for compactness, but the equations must be viewed as locally valid at a particular altitude. Moreover, the impact ionization term \( G \) depends on electron and neutral density, as well as electron temperature; and the coefficient of dissociative attachment \( \alpha_D \) depends on electron temperature.

The altitude profile of ion-ion recombination coefficient, \( \alpha(z) \), is obtained from Rosen and Hofmann [1981b] and illustrated in Figure 2.10a. Aerosol density, \( N_A(z) \) is obtained from Gringel
Figure 2.10: Altitude profiles of parameters employed in atmospheric chemistry model. (a) Ion-ion recombination coefficient [Rosen and Hofmann, 1981b]. (b) Aerosol density [Gringel et al., 1986, p. 169]. (c) Water vapor mixing ratio, in parts per million, mass [Granham et al., 1985].
et al. [1986, p. 169] and illustrated in Figure 2.10b. Finally, the water vapor mixing ratio, \( r_w \), is obtained from Grantham et al. [1985] and illustrated in Figure 2.10c.

In order to assess the relative importance of the many competing processes described in Tables 2.3 and 2.4, we will compare the magnitudes of the associated terms in (2.110). We compute the terms using the equilibrium densities at a particular altitude, and list them in Table 2.5. The densities employed are taken from Figure 2.11, which is obtained from the equilibrium solution.

<table>
<thead>
<tr>
<th>Particle</th>
<th>Process</th>
<th>Term</th>
<th>Rate (per m(^3) per s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>electrons</td>
<td>environmental ionization</td>
<td>( I )</td>
<td>0.98E+08</td>
</tr>
<tr>
<td></td>
<td>impact ionization</td>
<td>( G )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>detachment</td>
<td>( \lambda n_e N )</td>
<td>0.66E+12</td>
</tr>
<tr>
<td></td>
<td>three-body attachment</td>
<td>(-\alpha_3 n_e (N_{O2})^2)</td>
<td>-0.66E+12</td>
</tr>
<tr>
<td></td>
<td>dissociative attachment</td>
<td>(-\alpha_D n_e N_{O2})</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>recombination with small positive ions</td>
<td>(-\alpha_E n_p n_e)</td>
<td>-0.10E+05</td>
</tr>
<tr>
<td></td>
<td>recombination with cluster ions</td>
<td>(-\alpha_{EC} n_p n_e)</td>
<td>-0.92E+05</td>
</tr>
<tr>
<td></td>
<td>aerosol attachment</td>
<td>(-\beta N_{A} n_e)</td>
<td>-0.12E+03</td>
</tr>
<tr>
<td>small positive ions</td>
<td>environmental ionization</td>
<td>( I )</td>
<td>0.98E+08</td>
</tr>
<tr>
<td></td>
<td>impact ionization</td>
<td>( G )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>dehydration</td>
<td>( \lambda_c N_{N2} n_c )</td>
<td>0.62E+13</td>
</tr>
<tr>
<td></td>
<td>hydration</td>
<td>(-B_c N_{N2} r_w n_p)</td>
<td>-0.62E+13</td>
</tr>
<tr>
<td></td>
<td>recombination with small negative ions</td>
<td>(-\alpha n_p n_m)</td>
<td>-0.46E+08</td>
</tr>
<tr>
<td></td>
<td>recombination with electrons</td>
<td>(-\alpha_E n_p n_e)</td>
<td>-0.10E+05</td>
</tr>
<tr>
<td></td>
<td>aerosol attachment</td>
<td>(-\beta N_A n_p)</td>
<td>-0.20E+06</td>
</tr>
<tr>
<td>small negative ions</td>
<td>three-body attachment</td>
<td>(\alpha_3 n_e (N_{O2})^2)</td>
<td>0.66E+12</td>
</tr>
<tr>
<td></td>
<td>dissociative attachment</td>
<td>(\alpha_D n_e N_{O2})</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>recombination with small positive ions</td>
<td>(-\alpha n_p n_m)</td>
<td>-0.46E+08</td>
</tr>
<tr>
<td></td>
<td>recombination with cluster ions</td>
<td>(-\alpha_m n_e)</td>
<td>-0.51E+08</td>
</tr>
<tr>
<td></td>
<td>electron detachment</td>
<td>(-\lambda n_e N)</td>
<td>-0.66E+12</td>
</tr>
<tr>
<td></td>
<td>aerosol attachment</td>
<td>(-\beta N_A n_m)</td>
<td>-0.43E+06</td>
</tr>
</tbody>
</table>

Table 2.5: Magnitudes of processes affecting the density of each class of particles. Terms are evaluated for equilibrium densities, temperatures, and rate coefficients that are characteristic of 20 km altitude. The equilibrium state is specified at the end of this section. Continued.


<table>
<thead>
<tr>
<th>cluster ions</th>
<th>hydration</th>
<th>$B_C N_{N_2} N_{r_p} n_p$</th>
<th>$0.62E+13$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dehydration</td>
<td>$-\lambda C N_{N_2} n_e$</td>
<td>-0.62E+13</td>
<td></td>
</tr>
<tr>
<td>recombination with small negative ions</td>
<td>$-\alpha n_m n_e$</td>
<td>-0.51E+08</td>
<td></td>
</tr>
<tr>
<td>recombination with electrons</td>
<td>$-\alpha_{EC} n_e n_e$</td>
<td>-0.92E+05</td>
<td></td>
</tr>
<tr>
<td>aerosol attachment</td>
<td>$-\beta N_A n_e$</td>
<td>-0.22E+06</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.5 continued: Magnitudes of processes affecting the density of each class of particles. Terms are evaluated for equilibrium densities, temperatures, and rate coefficients that are characteristic of 20 km altitude. The equilibrium state is specified at the end of this section.

We conclude from Table 2.5 that the equilibrium electron and small negative ion densities are determined by a balance between three-body attachment of electrons to oxygen and the detachment of electrons from small negative ions. Similarly, the processes involving small positive ions and cluster ions are dominated by the hydration-dehydration reaction pair. We show results only for an altitude of 20 km, but the processes identified as dominant at equilibrium remain so throughout the simulation volume, although their magnitude does vary.

We note that the electron and small positive ion equations include a steady state production term, $I$. These term arises from the ionization of ambient neutrals by environmental sources. Since we are focusing on nighttime conditions, the principal environmental ionization source is the galactic cosmic ray flux. $I$ is a fixed function of altitude, resulting principally from the altitude variation of the neutral density. In the practice adopted here, the steady state production is determined at each point in the simulation volume so as to precisely balance the steady state loss represented by the remainder of the terms on the right hand sides of (2.110). In other words, $I$ is determined from the equilibrium conditions:

\[
\begin{align*}
\frac{d}{dt} n_e(t = -t_e) &= 0 \\
\frac{d}{dt} n_p(t = -t_e) &= 0 \\
\frac{d}{dt} n_m(t = -t_e) &= 0 \\
\frac{d}{dt} n_c(t = -t_e) &= 0
\end{align*}
\]

(2.111a)

(2.111b)

(2.111c)

(2.111d)

where $t = -t_e$ is the time immediately prior to the start of the simulation. In this manner, we ensure that the initial condition of the simulation is a steady state solution of the coupled rate equations.
(2.110). This would not be the case otherwise, since the initial density profile of the small positive ions obtained in (2.11) is an ad hoc representation of real conditions, and is not determined self-consistently from the ion chemistry formulation.

It is, of course, necessary and appropriate that the initial condition of the simulation be an equilibrium solution of both the chemical rate equations (2.110) and the system of fluid equations to be described in Section 2.7. This is because we envision the system as having existed in its initial state for an indeterminate—in principle infinite—length of time prior to the start of cloud electrification. On this basis, the initial state is by definition an equilibrium. Moreover, it is our goal to determine the response of the system to the stimulus provided by cloud electrification and lightning discharge. If the simulation were not to start from an equilibrium, we would observe not only its response to changing cloud charge, but also its response to pre-existing, non-equilibrium forces. Under such circumstances, the simulation would not represent a properly controlled experiment.

Our approach to defining the initial density and environmental ionization profiles begins with the given altitude profile of small positive ion density shown in Figure 2.3. As discussed in Section 2.2.2, this realization of \(n_p(z)\) forms the empirically justified ground truth upon which we build the equilibrium solution. Taking \(n_p\) as given, we then treat (2.111) as four equations in the unknowns \(I, n_e, n_m,\) and \(n_c\). These equations are solved locally, given the neutral densities at each altitude of interest.

To solve for the equilibrium state, we begin by combining (2.110) and (2.111), and rewrite the equilibrium conditions as:

\[
0 = I + \left[ g_i - g_o - \alpha_3 (N_{O2})^2 - \beta N_A \right] n_e + \lambda \dot{N} n_m - \alpha E n_p n_e - \alpha EC n_e n_e
\]  
(2.112a)

\[
0 = I - \left[ B_C N_{N2} N r_w + \beta N_A \right] n_p + g_i n_e - \alpha n_p n_m - \alpha E n_p n_e + \lambda_C N_{N2} n_e
\]  
(2.112b)

\[
0 = \left[ \alpha_3 (N_{O2})^2 + g_o \right] n_e - \alpha n_m n_p + \left[ \lambda N + \beta N_A \right] n_m
\]  
(2.112c)

\[
0 = B_C N_{N2} N r_w n_p - \lambda_C N_{N2} n_c - \alpha n_m n_c - \alpha EC n_e n_c - \beta N_A n_c
\]  
(2.112d)

where \(n_p(z)\) is given, and \(I(z), n_e(z), n_m(z),\) and \(n_c(z)\) are to be computed at a particular altitude. In addition, we have replaced the impact ionization term \(G\) by \(g_p n_e\), and the dissociative attachment term \(\alpha_{O2} n_e n_e\) by \(g_{oa} n_e\), where \(g_i\) and \(g_o\) are functions of both electron temperature and altitude.
We first note that by adding all of the equations (2.112a) through (2.112d) we obtain:

$$0 = -\beta N_A \left( n_{p0} + n_{e0} - n_{e0} - n_{m0} \right)$$  \hspace{1cm} (2.113)

Since \( \beta N_A \) is a known, non-zero coefficient, we conclude that in the steady state:

$$n_{e0} = n_{p0} + n_{e0} - n_{m0}$$  \hspace{1cm} (2.114)

Because all of the species under consideration are singly ionized, (2.114) is a statement of equilibrium charge neutrality. We will use (2.114) in place of (2.112a). Moreover, we will solve for the equilibrium ionization source \( I \) directly from (2.112b). This leaves us to solve (2.112c) and (2.112d) for \( n_{m0} \) and \( n_{e0} \). In order to simplify the presentation of these two equations, we introduce the following constants (2.115).

$$a_1 = \alpha_3 (N_{O2})^2 + g_a \hspace{1cm} a_3 = \lambda_c N_{N2}$$
$$a_2 = \alpha n_{p0} + \lambda N \hspace{1cm} b_1 = B_c N_{N2} N_{W}$$
$$a_3 = \beta N_A \hspace{1cm} b_2 = \alpha_{EC}$$
$$a_4 = \alpha$$

The equilibrium is then specified by:

$$n_{e0} = n_{p0} + n_{e0} - n_{m0}$$  \hspace{1cm} (2.116a)

$$I = (b_1 + a_3)n_{p0} - (g_i - \alpha_E n_{p0})n_{e0} + a_4 n_{p0} n_{m0} - a_3 n_{e0}$$  \hspace{1cm} (2.116b)

$$0 = a_1 n_{e0} - (a_2 + a_3)n_{m0} - a_4 n_{m0} n_{e0}$$  \hspace{1cm} (2.116c)

$$0 = b_1 n_{p0} - (a_3 + a_5)n_{e0} - a_4 n_{m0} n_{e0} - b_2 n_{e0} n_{e0}$$  \hspace{1cm} (2.116d)

Finally, we manipulate (2.116a), (2.116c) and (2.116d) to obtain the ultimate equations:

$$n_{e0} = n_{p0} + n_{e0} - n_{m0}$$  \hspace{1cm} (2.117a)

$$I = (b_1 + a_3)n_{p0} - (g_i - \alpha_E n_{p0})n_{e0} + a_4 n_{p0} n_{m0} - a_3 n_{e0}$$  \hspace{1cm} (2.117b)

$$n_{m0} = \frac{a_1 (n_{e0} + n_{p0})}{a_1 + a_2 + a_3 + a_4 n_{e0}}$$  \hspace{1cm} (2.117c)

$$0 = n_{e0}^3 + \left[ \frac{a_2 + a_3}{a_4} + \frac{a_1 + a_3 + a_5 + b_2 n_{p0}}{b_2} \right] n_{e0}^2 + \left[ \frac{a_4 n_{p0} (a_1 - b_1)}{a_4 b_2} + \frac{(a_1 + a_2 + a_5)(a_3 + a_5) + (a_2 + a_3) b_2 n_{p0}}{a_4 b_2} \right] n_{e0} - \frac{(a_1 + a_2 + a_3) b_2 n_{p0}}{a_4 b_2}$$  \hspace{1cm} (2.117d)
where the root of the cubic equation (2.117d) is chosen to ensure strictly positive \( n_{e0} \).

The procedure for obtaining a set of equilibrium values \((n_{e0}, n_{m0}, n_{e0}, I)\) appropriate to a given \( n_{po} \), altitude, and electron temperature is as follows. First, (2.117d) is solved for \( n_{e0} \) [Press et al., 1994, p. 178]. The resulting positive root is then used in (2.117c) to obtain \( n_{m0} \). Next, (2.117a) is used to find \( n_{e0} \), and finally (2.117b) provides \( I \). Figure 2.11 presents the equilibrium densities and environmental ionization rates obtained in this manner.

We can validate the steady state production term obtained from (2.117b) by comparing it to observed quiescent atmospheric ionization rates. Typical, background ionization rates presented by Reid [1986] show a rate of 1.0E+06 per m\(^3\) per s due to galactic cosmic rays at an altitude of 30 km. This is quite consistent with our calculated results shown in Figure 2.11b.

### 2.7 Fluid Model

In our discussion of the electron distribution function in Section 2.5.2, we introduced the Boltzmann Equation [Kraus and Trivelpiece, 1986, p. 311]:

\[
\frac{\partial}{\partial t} f_{\alpha} + \vec{v} \cdot \vec{\nabla} f_{\alpha} - \frac{e}{m} (\vec{E} + \vec{v} \times \vec{B}) \cdot \frac{\partial}{\partial \vec{v}} f_{\alpha} = \frac{\partial f_{\alpha}}{\partial t} |_{\epsilon}
\]

(2.118)

where \( f_{\alpha} \) now represents the distribution function of species \( \alpha \). In principal, (2.118) could be solved for \( f_{\alpha} \) in the presence of known electric and magnetic fields, or solved together with Maxwell's Equations for both \( f_{\alpha} \) and the fields. This is the province of kinetic theory. As an alternative, we may take moments of (2.118) to obtain equations for the various velocity moments of \( f_{\alpha} \), including density, flux, and pressure. Strictly speaking, only by solving for moments of all orders would we obtain all of the information contained in the distribution function itself. However, as a practical matter, the lowest order moments are of primary interest, and we truncate the chain of moment equations at some low order, in our case: 2. It is then necessary to close the system of equations by postulating an equation of state to relate the highest order moment of \( f_{\alpha} \) contained in the system to the lower order moments. This alternative approach is known as fluid theory.

The kinetic and fluid approaches ultimately provide equivalent descriptions of reality, since the kinetic solution could be integrated after the fact to provide the fluid moments, and the fluid solution could likewise be used to reconstruct the distribution function. Moreover, neither the fluid approach nor the kinetic approach places any inherent constraints on the distribution function.
Figure 2.11: Equilibrium quantities from atmospheric chemistry model. (a) Steady state electron density, $n_e$, small positive ion density, $n_p$, small negative ion density, $n_m$, and total cluster ion density, $n_c$. (b) Steady state environmental ionization, $I$, obtained from equilibrium solution.
Since our examination of lightning effects on the lower ionosphere is preoccupied with particle flows, densities, and temperatures—in other words, the macroscopic properties of the medium—the fluid approach is better suited to our purposes, and we employ it in the simulation. We are thereby faced with solving the three lowest order moment equations: continuity, momentum, and energy transport. We will now consider each, in turn.

2.7.1 The Continuity Equation

The continuity equation is the zeroth order moment of the Boltzmann Equation. When we operate on (2.118) with \( \bar{n}_a \int d^3\vec{v} \), where the integral is taken over all of velocity space, and \( \bar{n}_a \) is a system average density for the species \( \alpha \), we obtain [Krahl and Trivelpiece, 1986, p. 85]:

\[
\frac{\partial n_a}{\partial t} + \nabla \cdot (n_a \vec{u}_a) = \int \frac{\partial f_a}{\partial t} \bigg|_c d^3\vec{v} \tag{2.119}
\]

where, consistent with our normalization of Section 2.5.1:

\[
n_a(\vec{r},t) = \bar{n}_a \int d^3\vec{v} f_a(\vec{r},\vec{v},t), \text{ and } n_a(\vec{r},t) \vec{u}_a(\vec{r},t) = \bar{n}_a \int d^3\vec{v} \vec{v} f_a(\vec{r},\vec{v},t) \tag{2.120}
\]

In the customary case, in which binary collisions do not change the density of species \( \alpha \), the right hand side of (2.119) is zero. However, in the present case we are including the effects of several processes that act as local particle sources. In particular, impact ionization, background ionization, and chemical interactions all enter into (2.110), which determines the local production or loss of each species. We identify the rate of increase in the density of species \( \alpha \) due to the chemical effects (2.110) with a production term, \( S_\alpha \). In addition, the cloud source of (2.16) augments the small positive ion population. The right hand side of (2.119) represents the rate of density increase of species \( \alpha \) due to all of these processes.

A separate continuity equation applies to each of the four charged particle species, electrons, small positive ions, small negative ions, and cluster ions.

\[
\frac{\partial n_e}{\partial t} + \nabla \cdot (n_e \vec{u}_e) = S_e \tag{2.121a}
\]

\[
\frac{\partial n_p}{\partial t} + \nabla \cdot (n_p \vec{u}_p) = S_p + \frac{1}{e} \frac{\partial \rho_s}{\partial t} \tag{2.121b}
\]

\[
\frac{\partial n_m}{\partial t} + \nabla \cdot (n_m \vec{u}_m) = S_m \tag{2.121c}
\]
\[
\frac{\partial n_e}{\partial t} + \mathbf{\nabla} \cdot (n_e \mathbf{\bar{u}}_e) = S_e
\]  

(2.121d)

As discussed in Section 2.2.2, the cluster ions are the least mobile of all the species and are, in fact, virtually immobile. As a result, (2.121d) simplifies to:

\[
\frac{\partial n_e}{\partial t} = S_e
\]  

(2.122)

and cluster ion density responds only to local production and loss. Since cluster ion velocity plays no role in determining cluster ion density, the dynamics of cluster ions will not be considered further. In fact, the source term \(S_e\) is precisely the right hand side of the chemical rate equation (2.110d), and the two equations (2.122) and (2.110d) are identical. The evolution of the cluster ion density is thus determined entirely by the chemical effects model of Section 2.6.

Continuity applies not only to the flow of particles, but also to the flow of electrical current. We can obtain the form of current continuity that is appropriate to the present circumstances by taking a sum of the particle continuity equations, weighted by the charge carried by the individual particles:

\[
\frac{\partial}{\partial t} \left[ \sum_a q_a n_a \right] + \mathbf{\nabla} \cdot \left[ \sum_a q_a n_a \mathbf{\bar{u}}_a \right] = \sum_a q_a S_a + \frac{1}{e} \frac{\partial \rho_s}{\partial t}
\]  

(2.123)

However, the bracketed terms are precisely the definitions of charge density \(\rho\), and current density \(\mathbf{j}\):

\[
\rho = \sum_a q_a n_a \quad \text{and} \quad \mathbf{j} = \sum_a q_a n_a \mathbf{\bar{u}}_a
\]  

(2.124)

so that current continuity may be written:

\[
\frac{\partial}{\partial t} \rho + \mathbf{\nabla} \cdot \mathbf{j} = \sum_a q_a S_a + \frac{1}{e} \frac{\partial \rho_s}{\partial t}
\]  

(2.125)

The four equations of particle continuity and the equation of current continuity are all consistent with one another. However, by virtue of (2.123) they are also redundant. As a result, we will solve continuity equations only for the three ion species and current density. The electron density can then be obtained from the definition of charge density in terms of the ion densities and the charge density:

\[
n_e = n_p + n_e - n_m - \frac{\rho}{e}
\]  

(2.126)
2.7.2 The Momentum Equation

The momentum equation is the first velocity moment of the Boltzmann Equation. We now operate on (2.118) with \( m_a \bar{n}_a \int d^3 \vec{v} \bar{v} \), and use the fact that \( f_a \) must approach 0 as \( \|\vec{v}\| \to \infty \) to obtain:

\[
m_a \frac{\partial}{\partial t}(n_a \bar{u}_a) + m_a \vec{\nabla} \cdot (n_a \vec{u}_a \bar{u}_a) = -\vec{\nabla} \cdot P_a + q_a n_a (\vec{E} + \vec{u}_a \times \vec{B}) + m_a \bar{n}_a \int d^3 \vec{v} \bar{v} \frac{\partial f_a}{\partial t} \Bigg|_c
\]

(2.127)

where \( q_a \) is the charge on species \( \alpha \), and the \((i,j)\) element of the pressure tensor \( P_a \) is given by:

\[
P_{ai,j} = m_a \bar{n}_a \int (v_i - u_i)(v_j - u_j) f_a d^3\vec{v}
\]

(2.128)

In order to evaluate the collision term, we need an explicit form for the Boltzmann collision integral. For this purpose we use the Krook collision model [Krall and Trivelpiece, 1986, p. 316]:

\[
\frac{\partial f_a}{\partial t} \bigg|_c = \frac{f_{a0} - f_a}{\tau_a}
\]

(2.129)

where \( f_{a0} \) is the equilibrium distribution function for species \( \alpha \), and \( \tau_a \) is the mean collision time for species \( \alpha \). The chief simplification here is the use of a velocity-independent mean collision time.

This is a well-founded approximation for the case of collisions with neutral molecules [Krall and Trivelpiece, 1986, p. 317] since the electron-neutral interaction involves extremely short-range forces. We have found electron-neutral collisions to be the dominant collision type in the weakly ionized medium under consideration (see Section 2.5.1). Both elastic and inelastic collisions cause electron momentum loss, and both processes are incorporated in the momentum transfer collisional cross sections introduced in Section 2.5.3. Using the mean collision frequency calculated from these cross sections according to (2.97), we write:

\[
\tau_e = \frac{1}{\nu_{mom}(T_e)}
\]

(2.130)

and corresponding results hold for the ionic species. Given \( \tau_e \), and taking \( f_{a0} \) to be isotropic, like the distribution of the background neutrals, we compute the integral in (2.127) to obtain:

\[
m_a \frac{\partial}{\partial t}(n_a \bar{u}_a) + m_a \vec{\nabla} \cdot (n_a \vec{u}_a \bar{u}_a) = -\vec{\nabla} \cdot P_a + q_a n_a (\vec{E} + \vec{u}_a \times \vec{B}) - \frac{m_a n_a \bar{u}_a}{\tau_a}
\]

(2.131)

We may rewrite this as:

\[
m_a n_a \frac{d \bar{u}_a}{dt} + m_a \bar{u}_a \left[ \frac{\partial n_a}{\partial t} + \vec{\nabla} \cdot (n_a \bar{u}_a) \right] = -\vec{\nabla} \cdot P_a + q_a n_a (\vec{E} + \vec{u}_a \times \vec{B}) - \frac{m_a n_a \bar{u}_a}{\tau_a}
\]

(2.132)
where we have introduced the co-moving derivative:

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \bar{u}_\alpha \cdot \vec{\nabla}$$

(2.133)

In the absence of an impact ionization source, the second term on the left-hand side of (2.132) would be zero by continuity, and we would recover the familiar form of the momentum equation. However, in the present case, we substitute (2.121) and find:

$$m_a n_a \frac{d \bar{u}_\alpha}{dt} = -\vec{\nabla} \cdot P_a + q_a n_a \left( \vec{E} + \bar{u}_\alpha \times \vec{B} \right) - \frac{m_a n_a \bar{u}_\alpha}{\tau} - m_a \bar{u}_\alpha S_a$$

(2.134)

Moreover, since we are dealing with an isotropic particle distribution (see Section 2.5.2), the pressure tensor becomes diagonal:

$$P_a = m_a n_a \int \bar{v} \bar{v} f_a d^3 \bar{v} = \frac{1}{3} m_a n_a I \int \bar{v}^2 f_a d^3 \bar{v} = I n_a \kappa T_a$$

(2.135)

where $I$ is a 3-by-3 unit matrix, and we have used the definition of effective temperature from (2.92). Using this result in (2.134), and recalling that the system is unmagnetized (see Section 2.1) our final form for the momentum equation becomes:

$$m_a n_a \frac{d \bar{u}_\alpha}{dt} = -\vec{\nabla} \cdot (n_a \kappa T_a) + q_a n_a \bar{E} - \frac{m_a n_a \bar{u}_\alpha}{\tau_a} - m_a \bar{u}_\alpha S_a$$

(2.136)

The mobility of species $\alpha$ may be obtained from (2.136). Of course, mobility is a well known parameter that emerges in identical form from either the fluid or the kinetic theory approach. Our purpose in deriving it from (2.136) is to call attention to the conditions under which the mobility is, strictly speaking, a meaningful concept. In particular, for a homogeneous medium in the steady state, the derivatives in (2.136) vanish, and we are left with:

$$0 = q_a n_a \bar{E} - \frac{m_a n_a \bar{u}_\alpha}{\tau_a} - m_a \bar{u}_\alpha S_a$$

(2.137)

Moreover, in the absence of particle production, this further reduces to:

$$0 = q_a n_a \bar{E} - \frac{m_a n_a \bar{u}_\alpha}{\tau_a}$$

(2.138)

which we may solve for:

$$\bar{u}_\alpha = \mu_a \bar{E}, \text{ where } \mu_a = \frac{q_a \tau_a}{m_a}$$

(2.139)

and $\mu_a$ is the mobility of species $\alpha$. In addition, we may compute the current density:
\[ \mathbf{J} = \sum_{a} q_a n_a \mathbf{u}_a = \sigma \mathbf{E} \] where \( \sigma = \sum_{a} q_a n_a \mu_a \) \hspace{1cm} (2.140)

and \( \sigma \) is the total conductivity of the medium. We note that (2.140) is a statement of Ohm's Law.

Thus, the mobility accurately relates the electric field to the drift velocity of species \( \alpha \) only under steady state conditions in a homogeneous medium with no particle production source. The same limitations apply to the use of the conductivity. In practice, the temporal variation in the system may be slow enough that a steady state is well approximated over the time scales of particle drift. Likewise, the spatial variation may be gradual enough that the medium can be treated as locally uniform. In both of these cases, mobility and conductivity are operationally useful concepts. Finally, whether or not mobility and conductivity are useful in the sense that Ohm's Law applies, they can still be used to characterize the medium.

### 2.7.3 The Energy Transport Equation

The energy transport equation is a second order velocity moment of the Boltzmann Equation, obtained by operating on (2.118) with \( \frac{1}{2} m_a \bar{n}_a \int d^3 \mathbf{v} v^2 \). If we define \( \bar{w} \) as the particle velocity relative to the mean drift, so that

\[ \bar{v} = \bar{w} + \bar{u}_a \quad \text{and} \quad \int \bar{w} f_a (\bar{w} + \bar{u}_a) d^3 \bar{w} = 0 \] \hspace{1cm} (2.141)

then the desired moment may be written:

\[
\frac{\partial}{\partial t} \left[ \frac{1}{2} n_a m_a u_a^2 + \frac{1}{2} \bar{n}_a m_a \int w^2 f_a d^3 \bar{w} \right] + \bar{\nabla} \cdot \left[ \left( \frac{1}{2} m_a v^2 \right) \bar{n}_a f_a \bar{v} d^3 \bar{v} \right] - n_a q_a \mathbf{E} \cdot \bar{u}_a
\]

\[
= -\frac{\partial}{\partial t} \int \frac{1}{2} \bar{n}_a m_a v^2 f_a d^3 \bar{v} \bigg|_{c}
\] \hspace{1cm} (2.142)

where the acceleration term has been integrated using the technique of integration by parts, together with the knowledge that \( f_a \) extinguishes the integrand at large \( v \). We recognize the first integral on the left hand side of (2.142) as the average kinetic energy density of the particles, which we identified with the effective temperature in (2.92) for the case of an isotropic distribution in the local approximation.

\[ \frac{1}{2} \bar{n}_a m_a \int w^2 f_a d^3 \bar{w} = \frac{1}{2} n_a \kappa T_a \] \hspace{1cm} (2.143)

This is, in fact, the general definition of effective temperature [Krall and Trivelpiece, 1986, p. 84]. The form of (2.92) is a special case that is obtained under the local approximation,
\[ \tilde{n}_a f_a (\tilde{\nu}, \tilde{\varphi}) = n_a (\tilde{\varphi}) f_a (\tilde{\nu}) \] (2.144)

The second integral on the left hand side of (2.142) is defined as the heat flux vector [Krall and Trivelpiece, 1986, p. 84], the flux of kinetic energy associated with particles of species \( \alpha \):

\[ \tilde{H}_\alpha = \int \left( \frac{1}{2} m_a \nu^2 \right) \tilde{n}_a f_a \nu d^3 \tilde{\nu} \] (2.145)

We can evaluate this integral by using the substitution (2.141) and the definition of the pressure tensor (2.128).

\[ \tilde{H}_\alpha = \frac{1}{2} n_a m_a u_a^2 \tilde{u}_a + \frac{3}{2} n_a \kappa T_a \tilde{u}_a + \tilde{u}_a \cdot P_a \] (2.146)

Using (2.135) for the isotropic pressure tensor, the heat flux vector becomes:

\[ \tilde{H}_\alpha = \frac{1}{2} n_a m_a u_a^2 \tilde{u}_a + \frac{5}{2} n_a \kappa T_a \tilde{u}_a \] (2.147)

Substituting (2.143) and (2.145) in (2.142), we obtain:

\[ \frac{\partial}{\partial t} \left[ \frac{1}{2} n_a m_a u_a^2 + \frac{3}{2} n_a \kappa T_a \right] + \tilde{\nabla} \cdot \tilde{H}_\alpha - n_a q_a \tilde{E} \cdot \tilde{u}_a = \left. \frac{\partial}{\partial t} \left[ \frac{1}{2} \tilde{n}_a m_a \nu^2 f_a d^3 \tilde{\nu} \right] \right|_c \] (2.148)

Using (2.147) in (2.148), we find

\[ \frac{\partial}{\partial t} \left[ \frac{1}{2} n_a m_a u_a^2 + \frac{3}{2} n_a \kappa T_a \right] + \tilde{\nabla} \cdot \left[ \frac{1}{2} n_a m_a u_a^2 \tilde{u}_a + \frac{5}{2} n_a \kappa T_a \tilde{u}_a \right] - n_a q_a \tilde{E} \cdot \tilde{u}_a \] (2.149)

Rearranging this result:

\[ \frac{1}{2} m_a n_a \frac{d u_a^2}{dt} + \frac{3}{2} m_a u_a^2 \left( \frac{\partial n_a}{\partial t} + \tilde{\nabla} \cdot (n_a \tilde{u}_a) \right) + \frac{3}{2} n_a \frac{d \kappa T_a}{dt} + \frac{3}{2} \tilde{\kappa} T_a \frac{\partial n_a}{\partial t} + \tilde{\nabla} \cdot (n_a \tilde{u}_a) \] (2.150)

The second and fourth terms on the left hand side of (2.150) would vanish in the absence of a particle production source, but when we substitute (2.121) we instead find:

\[ n_a \frac{d}{dt} \left( \frac{1}{2} m_a u_a^2 + \frac{3}{2} \kappa T_a \right) + \left( \frac{1}{2} m_a u_a^2 + \frac{3}{2} \kappa T_a \right) S_n + \tilde{\nabla} \cdot (n_a \kappa T_a \tilde{u}_a) - n_a q_a \tilde{E} \cdot \tilde{u}_a = \left. \frac{\partial}{\partial t} \left[ \frac{1}{2} \tilde{n}_a m_a \nu^2 f_a d^3 \tilde{\nu} \right] \right|_c \] (2.151)

It remains to address the collisional energy loss term on the right hand side of (2.151). Computing this collision integral forces us to specify an approximate, assumed distribution function for particles of species \( \alpha \). This same act also closes our set of fluid equations by
completely specifying all higher order interactions between the particles and their neutral background.

The collisional energy loss is composed of two parts, one part given up directly to the kinetic (thermal) energy of the neutrals, which we refer to as elastic loss, and the other part given up to the potential energy of the neutrals, in the form of molecular and electronic excitation, which we refer to as inelastic loss. We write the elastic loss according to a model by Gurevich [1978, p. 19]:

$$\frac{\delta}{\tau_a} (\kappa T_a - \kappa T_N)$$

(2.152)

where $\delta$ is the fraction of particle kinetic energy transferred in a single collision, $T_N$ is the neutral temperature, and $\tau_a$ is the mean momentum transfer collision time for particles of species $\alpha$. This result assumes Maxwellian populations, which is accurate for the neutrals. For the purposes of calculating energy loss, this assumption is a good approximation for electrons as well, as discussed in Section 2.5.3. The value of $\delta$ is obtained as a straightforward kinematic result [Wadzinski and Jasperse, 1982, p. 46]. For cross sections that are independent of scattering angle, as are the momentum transfer cross sections, we have:

$$\delta = \frac{2m}{M}$$

(2.153)

where $m$ is the mass of the less massive particle, $M$ is the mass of the other, and we assume $m \ll M$. Needless to say, using (2.153) under this assumption limits consideration to electron collisions with neutrals. The formulation of the energy transport equation to this point has otherwise been general. The value of $\delta$ for ions is of order unity. Fortunately, we will see in Section 2.7.5 that order of magnitude knowledge of $\delta$ is adequate to determine ion temperature.

By construction, (2.152) is simply the rate at which kinetic energy is transferred from a single particle of species $\alpha$ to the kinetic energy of the neutral population, minus the rate of transfer in the reverse direction. We will write the inelastic loss per particle as a general function of species $\alpha$ temperature, $\Phi(T_a)$. We have specified the actual form of this function in the case of electrons in (2.99). This allows us to rewrite (2.151) in its final form:

$$n_a \frac{d}{dt} \left( \frac{1}{2} m_a u_a^2 + \frac{1}{2} \kappa T_a \right)$$

$$= -\left( \frac{1}{2} m_a u_a^2 + \frac{1}{2} \kappa T_a \right) S_a - \nabla \cdot (n_a \kappa T_a \bar{u}_a) + n_a q_a \bar{E} \cdot \bar{u}_a - \frac{\delta}{\tau_a} n_a \kappa (T_a - T) - n_a \Phi(T_a)$$

(2.154)
2.7.4. Normalization of the Fluid Equations

We simultaneously nondimensionalize and normalize the continuity, momentum, and energy transport equations. We accomplish this by rewriting all of the variables in the problem, both dependent and independent, in terms of normalized variables that have been rescaled by representative values of the original, dimensional variables. Each of the normalized variables is identified by a tilde. We define normalizing parameters as follows:

\[
\begin{align*}
    t &= \tau_0 \tilde{t} \\
    x &= L \tilde{x} \\
    u_a &= \mu a_0 \tilde{E}_0 \tilde{u}_a \\
    E &= \tilde{E}_0 \tilde{E} \\
    T_a &= \tau a_0 \tilde{T}_a \\
    S_a &= \tilde{S}_a \frac{a_0}{\tau_0} \\
    n_a &= \tilde{n}_a \\
    q_a &= e \tilde{q}_a \\
    \frac{\partial \rho_s}{\partial t} &= \frac{e n_{p_0} \rho_s}{\tau_0} \\
\end{align*}
\]

(2.155)

where the normalizing factors apply to each component of a vector, and \( x \) is used to represent any spatial independent variable. Note that the net particle production is simply nondimensionalized, and not normalized, in this process. We use the scale factors identified in (2.155) to form the four dimensionless parameters listed in Table 2.6.

<table>
<thead>
<tr>
<th>Dimensionless Parameter</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = \frac{\tau_0 \mu a_0 E_0}{L} )</td>
<td>measure of drift distance versus feature size</td>
</tr>
<tr>
<td>( \beta = \frac{\kappa T a_0}{e E_0 L} )</td>
<td>measure of particle thermal energy versus work done by particle against electric field</td>
</tr>
<tr>
<td>( \eta = \frac{\kappa T a_0}{m_a \mu a_0 E_0^2} )</td>
<td>measure of thermal energy versus drift kinetic energy</td>
</tr>
<tr>
<td>( \frac{\tau_0}{\tau_a} )</td>
<td>measure of system time scale versus mean momentum transfer collision time</td>
</tr>
</tbody>
</table>

Table 2.6: Definitions of dimensionless parameters employed in fluid formulation.

These dimensionless parameters allow us to rewrite the continuity, momentum, and energy transfer equations. We first examine the co-moving derivative, which appears in two of the original equations, and is defined in (2.133).
\[
\frac{d}{d\tilde{t}} = \frac{\partial}{\partial\tilde{t}} + \alpha \tilde{u} \cdot \nabla
\] (2.156)

The continuity equations (2.121) become:

\[
\frac{\partial \tilde{n}_a}{\partial\tilde{t}} + \alpha \nabla \cdot (\tilde{n}_a \tilde{u}_a) = \tilde{S}_a \quad \text{for index } \alpha \neq p
\] (2.157a)

\[
\frac{\partial \tilde{n}_p}{\partial\tilde{t}} + \alpha \nabla \cdot (\tilde{n}_p \tilde{u}_p) = \tilde{S}_p + \frac{\partial \tilde{n}_a}{\partial\tilde{t}}
\] (2.157b)

Similarly, we obtain normalized versions of the momentum equation (2.136):

\[
\tilde{n}_a \frac{d}{d\tilde{t}} \tilde{u}_a = -\beta \frac{\tau_0}{\tau_a} \tilde{E} \left( \tilde{n}_a \tilde{T}_a \right) + \frac{\tau_0}{\tau_a} \tilde{q}_a \tilde{n}_a \tilde{E} - \frac{\tau_0}{\tau_a} \tilde{n}_a \tilde{u}_a - \tilde{u}_a \tilde{S}_a
\] (2.158)

and the energy transport equation (2.154):

\[
\tilde{n}_a \frac{d}{d\tilde{t}} \left( \frac{1}{2} \tilde{u}_a^2 + \frac{1}{2} \eta \tilde{T}_a \right) + \tilde{S}_a \left( \frac{1}{2} \tilde{n}_a \tilde{T}_a + \frac{1}{2} \eta \tilde{T}_a \right) + \beta \frac{\tau_0}{\tau_a} \tilde{E} \cdot \nabla \left( \tilde{n}_a \tilde{u}_a \tilde{T}_a \right) = \frac{\tau_0}{\tau_a} \tilde{q}_a \tilde{n}_a \tilde{E} \cdot \tilde{u}_a
\]

(2.159)

\[
= -\frac{1}{2} \delta \frac{\tau_0}{\tau_a} \tilde{n}_a \left( \tilde{T}_a - \frac{T_N}{T_{a0}} \right) - \frac{\tau_0}{\kappa T_{a0}} \eta \tilde{n}_a \Phi
\]

Note once again that the production source, \( \tilde{S}_a \), is not normalized, merely nondimensionalized, and the same is true of the final term in (2.159), representing inelastic energy loss.

2.7.5 Steady State Fluid Results

The variables, both dependent and independent, in properly normalized equations such as (2.157), (2.158) and (2.159), are all of order unity. This serves two purposes. First, it permits us to compare the magnitudes of the several terms on the basis of their coefficients. Second, it makes the equations more tractable in numerical solution schemes. We will now make use of the normalization to estimate the magnitudes of the terms in (2.158) and (2.159). No estimate may be made of those few terms that we have left unnormalized.

In order to evaluate the dimensionless parameters of Table 2.6, we must provide representative values for the many scale factors. We base our system evolution time scale, \( \tau_0 \), and feature scale size, \( L \), on observations of sprites and related activity, as summarized in Section 1.1.1. Since sprite duration is several milliseconds, we choose a time scale of 1 ms to resolve the most rapidly evolving features. A typical linear dimension of a sprite feature is 10 km, while the scale height of charged particle density variation is several kilometers, so we choose \( L \) to be 1 km.

102
Choosing somewhat low values for these system-wide parameters is the conservative approach, in the sense that we thereby anticipate resolving the smallest spatial, and shortest temporal, features.

The remaining parameters in (2.155), namely \( E_0 \), \( n_{ao} \), \( \mu_{ao} \), and \( T_{ao} \), must be chosen appropriately for each species and altitude in order to obtain a correct normalization. We consider two representative altitudes, 30 km and 70 km. Representative electric field changes at 30 km and 70 km altitude over a thunderstorm are on the order of 100 V/m and 1 V/m, respectively, based upon a typical measured value of 1 kV/m taken at 20 km over an active storm cell [Blakeslee, et al., 1989], and extrapolated to higher altitude assuming a monopole cloud charge at 10 km altitude over a conductive ground plane. Characteristic densities, \( n_{ao} \), are taken to be the equilibrium, quiescent densities of Figure 2.11. We use the mildly heated value of 0.1 eV as the representative temperature for all species based upon the geometric mean of the neutral background temperature and the elevated electron temperature in a sprite, as deduced by Green et al. [1995]. Finally, particle mobilities are calculated from (2.139) using the appropriate momentum transfer collision rates, ion mass of NO\(^+\) (5.0E-26 kg), and the assumed temperature. The resulting values of the dimensionless parameters are listed in Table 2.7.

<table>
<thead>
<tr>
<th></th>
<th>Electrons</th>
<th>Ions</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_0 / \tau_e )</td>
<td>2.0E+06</td>
<td>1.0E+04</td>
</tr>
<tr>
<td>( \eta )</td>
<td>7.0E+02</td>
<td>4.4E+01</td>
</tr>
<tr>
<td>( \beta )</td>
<td>1.0E-06</td>
<td>1.0E-06</td>
</tr>
<tr>
<td>( \mu_{ao} ) (m(^2)/Vs)</td>
<td>5.0E+01</td>
<td>2.0E+04</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>5.0E-03</td>
<td>2.0E-02</td>
</tr>
<tr>
<td>( \tau_0 ) (s)</td>
<td>5.0E-10</td>
<td>1.0E-07</td>
</tr>
<tr>
<td></td>
<td>70 km altitude</td>
<td>30 km altitude</td>
</tr>
<tr>
<td></td>
<td></td>
<td>70 km altitude</td>
</tr>
<tr>
<td></td>
<td>30 km altitude</td>
<td>70 km altitude</td>
</tr>
<tr>
<td></td>
<td>1.2E-07</td>
<td>2.4E-05</td>
</tr>
<tr>
<td></td>
<td>8.5E+01</td>
<td>2.1E-01</td>
</tr>
<tr>
<td></td>
<td>8.5E+05</td>
<td>2.1E-05</td>
</tr>
</tbody>
</table>

Table 2.7: Values of the dimensionless parameters used in the fluid model.

We may draw several conclusions from the parameter values shown above. First, since \( \alpha \) is much smaller for ions than for electrons, (2.156) suggests that ion drifts are much less important than electron drifts. Next, since \( \tau_0 / \tau_e \gg 1 \), the inertial term in (2.158) is negligible compared to the force terms in the case of electrons, and we may solve the electron dynamics in the steady state:

\[
0 = -\beta \nabla (\tilde{n}_e \tilde{T}_e) + \tilde{q}_e \tilde{n}_e \tilde{E} - \tilde{\eta}_e \tilde{u}_e - \frac{\tau_e}{\tau_0} \tilde{u}_e S_e
\]  

(2.160)
Moreover, since $\beta << 1$ as well, it turns out that Ohm's Law is a good approximation as well, in cases for which electron production is not significant. The same approximations are less adequate for ions, particularly at high altitude. Thus, although Ohm's Law could be reasonably applied to ions in most of the simulation volume, we will continue use the full momentum equation (2.158) in their case.

The energy transport equation (2.159) may be simplified considerably in view of the parameter values in Table 2.7. In the first place, since $\eta >> 1$ we can ignore the drift kinetic energy compared to the thermal energy for each species. This is consistent with our neglecting drifts when we discussed the isotropy of the electron velocity distribution in Section 2.5.2. We are thereby permitted to rewrite (2.159) as:

$$\frac{\tau_0}{\tau_a} \eta \vec{E} \cdot \vec{u}_a - \beta \frac{\tau_0}{\tau_a} \vec{v} \cdot (\vec{u}_a \vec{u}_a \vec{E}) - \left( \frac{3}{2} \eta \vec{m}_a \right) \vec{S}_a - \frac{2}{3} \delta \eta \frac{\tau_0}{\tau_a} \vec{n}_a \left( \frac{\vec{T}_a - T_{i0}}{T_{a0}} \right) - \frac{\tau_0}{\kappa} \eta \vec{n}_a \Phi$$

(2.161)

Next, we recall that while $\delta$ is extremely small for electrons ($3.9E-05$ for collisions with molecular nitrogen) it is actually of order unity for ions. As a result, the elastic relaxation term, which contains the product of the two large parameters $\eta$ and $\tau_0/\tau_a$, overwhelms the remaining terms in the ion energy transport equation. If this is the case, then it must be true that

$$\frac{\left| \vec{T}_i - T_{i0} \right|}{T_{i0}} << 1$$

(2.162)

and ion temperature must remain extremely close to the neutral background temperature. Strictly speaking, we cannot neglect the impact ionization source and inelastic loss terms in (2.161), because they are not normalized and we cannot speculate in general regarding their magnitude. However, (2.162) is nevertheless an accurate statement, because both impact ionization and inelastic loss are essentially zero when $T_i \approx T_i$.

By contrast, elastic thermal relaxation is vastly slower for electrons due to the small value of $\delta$ that applies to them. Their temperature is therefore able to depart significantly from that of the neutral background. We are able to determine the electron temperature from the steady state form of (2.161) because $\tau_0/\tau_e >> \eta$ for electrons and the derivative term may be ignored.
0 = e\vec{n}_e \vec{E} \cdot \vec{u}_e + \beta e \vec{N} \cdot (\vec{n}_i \vec{u}_i \vec{T}_i) + \left(\frac{1}{2} \eta \vec{T}_e + \frac{1}{2} \delta \eta \vec{n}_i \left(\frac{\vec{T}_e}{T_{e0}} - \frac{T_i}{T_{e0}}\right) + \frac{\tau_e}{\kappa T_{e0}} \eta \vec{n}_i \Phi \right) \tag{2.163}

We conclude this chapter by compiling the complete list of fluid equations that apply to the system under consideration. These consist of continuity equations for electrons and for ions, a steady state momentum equation for electrons, a time-varying momentum equation for ions, a steady state energy transport equation for electrons, and a constant temperature equation for ions. The index \(i\) used for the ion quantities represents each of the ion species separately. For the purposes of actually solving the equations, we use a value of reference temperature that is the same for all particles, namely the neutral background temperature, \(T_{e0} = T_{i0} = T_{e_f}\). We will also make explicit the dependence of parameters \(\alpha\) and \(\beta\) on species, while \(\eta\) is now applied only to electrons.

\[
\frac{\partial \vec{n}_e}{\partial t} + \alpha_e \vec{v} \cdot (\vec{n}_e \vec{u}_e) = \vec{S}_e \tag{2.164a}
\]

\[
\frac{\partial \vec{n}_p}{\partial t} + \alpha_p \vec{v} \cdot (\vec{n}_p \vec{u}_p) = \vec{S}_p + \frac{\delta \vec{P}}{\partial t} \tag{2.164b}
\]

\[
\frac{\partial \vec{n}_m}{\partial t} + \alpha_m \vec{v} \cdot (\vec{n}_m \vec{u}_m) = \vec{S}_m \tag{2.164c}
\]

\[
\frac{\partial \vec{n}_i}{\partial t} = \vec{S}_i \tag{2.164d}
\]

\[
0 = \beta_e \vec{v} (\vec{n}_e \vec{T}_e) + e \vec{n}_e \vec{E} + \vec{n}_e \vec{u}_e + \frac{\tau_e}{\tau_{e0}} \vec{u}_e \vec{S}_e \tag{2.164e}
\]

\[
\vec{n}_i \frac{d \vec{u}_i}{dt} = -\beta_i \frac{\tau_{i0}}{\tau_i} \vec{v} (\vec{n}_i \vec{T}_i) + \frac{\tau_{0}}{\tau_i} e \vec{n}_i \vec{E} - \frac{\tau_{0}}{\tau_i} \vec{n}_i \vec{u}_i - \vec{u}_i \vec{S}_i \tag{2.164f}
\]

\[
0 = e\vec{n}_e \vec{E} \cdot \vec{u}_e + \beta_e \vec{v} \cdot (\vec{n}_e \vec{u}_e \vec{T}_e) + \left(\frac{1}{2} \eta \vec{T}_e + \frac{1}{2} \delta \eta \vec{n}_i \vec{T}_i - 1\right) + \frac{\tau_e}{\kappa T_{e0}} \eta \vec{n}_i \Phi \tag{2.164g}
\]

\[
\vec{T}_i = 1 \tag{2.164h}
\]

105
Chapter 3

Numerical Methods

Although the numerical techniques employed in this effort are well established, some discussion of them is warranted. First, it is necessary to explain the novel manner in which individual numerical models have been combined into a comprehensive simulation. Second, we need to motivate the choice of the particular numerical approaches used. Finally, we must verify the stability and accuracy of the techniques.

In this chapter, we will provide an overview of the solution process, followed by a discussion of the three major numerical methods employed: the flux-corrected transport (FCT) algorithm, the finite differences cylindrical Poisson solver, and the Runge-Kutta method. We include an examination of the accuracy and stability of each of the methods.

3.1 Overview of the Solution Process

The fluid model equations (2.164), which were the culmination of Chapter 2, govern the nondimensionalized forms of the temperatures, \( T_a \); velocities, \( \vec{u}_a \); and densities, \( n_a \); of all the mobile charged species in the simulation. The determining quantities that enter into this system of equations, namely the electric field, \( \vec{E} \); particle production, \( S_a \); mean collision times, \( \tau_a \); as well as the initial values for the dependent variables; are all provided by supporting models specified in Chapter 2.

Our fundamental, underlying approach is finite differences. We begin by making the continuous spatial and temporal domain of the problem into a grid of discrete points at which we evaluate and track the dependent variables. Derivatives in time and space are approximated by differences of the variables on this grid, thereby transforming the partial differential equations of interest into a system of coupled algebraic equations.

First and foremost, a finite differences technique is applied to the particle continuity equations, whereby knowledge of particle velocities at one time step determines the evolution of the particle densities from that time step to the next. The array of particle densities everywhere on the
grid constitutes the ultimate state variable of the system. As a practical matter, all of the other supporting variables must be discretized on the same grid and serve to determine the velocities, which then determine the evolution of the densities.

By taking advantage of the cylindrical symmetry of the system, we can represent our three dimensional simulation volume by using a two dimensional grid. We divide the radial coordinate, \( r \), of the system into 64 discrete segments, spanning radii from 0 km to 64 km. The vertical coordinate, \( z \), is divided into 65 discrete segments from 0 km to 90 km altitude. The particular choice of 65 points versus 64 is inspired by efficiency considerations for the Poisson solver.

The grid positions themselves represent the centers of individual grid cells. They are expressed as \( (r_i, z_k) \). In the general case, which includes unevenly spaced grid positions, we denote the positions of the cell boundaries by \( (r_{i+\frac{1}{2}}, z_{k+\frac{1}{2}}) \), and locate them midway between successive grid positions:

\[
\begin{align*}
  r_{i+\frac{1}{2}} &= \frac{1}{2}(r_i + r_{i+1}) \quad \text{and} \quad z_{k+\frac{1}{2}} &= \frac{1}{2}(z_k + z_{k+1})
\end{align*}
\]  

(3.1)

The grid spacings, \( \Delta r \) and \( \Delta z \), are then defined to be the distance between the cell boundaries:

\[
\begin{align*}
  \Delta r_i &= r_{i+\frac{1}{2}} - r_{i-\frac{1}{2}} = \frac{1}{2}(r_{i+1} - r_i) \quad \text{and} \quad \Delta z_k &= z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}} = \frac{1}{2}(z_{k+1} - z_k)
\end{align*}
\]  

(3.2)

While uniform spacing of the grid points in the radial coordinate is natural based upon the uniformity of ambient conditions in the horizontal direction, several approaches to the spacing of grid points along the \( z \)-axis were examined. The purpose of this examination was to optimize vertical resolution as a function of altitude for a fixed number of grid points, in light of the fact that electron and small ion density vary most rapidly with altitude above 60 km. However, no significant advantage, either in feature resolution or in economy of grid points, was obtained by using a nonlinear grid. In fact, grid points spaced linearly in the vertical dimension are well-suited to the problem at hand. This is because particle mobilities and drift velocities increase rapidly with altitude, due to the rapid decline of neutral density. As we will see in Section 3.2, small grid spacings can degrade the stability of our finite differences solution in the presence of high drift velocities. Thus, a nonlinear grid spaced widely enough to provide acceptable stability at high altitudes yields unacceptably poor grid resolution at low altitudes, for any reasonable number of grid points. A linear grid with sufficient high altitude resolution works well also at low altitudes and is thus an effective compromise.

The grid spacing is chosen to be smaller than the minimum high altitude scale length of 2.5 km (see Section 2.2.2), and consistent with the 1 km scale length of the fluid formulation (see
Section 2.7.4). A linear grid of 64 cells in the radial dimension, and 65 cells in the vertical dimension yields a uniform grid spacing of 1 km in \( r \), and 1.38 km in \( z \), so that:

\[
\begin{align*}
\Delta r_i &= \Delta r = 1 \text{ km} \\
r_i &= (i - \frac{1}{2}) \Delta r \\
r_{i+\frac{1}{2}} &= i \Delta r \\
r_{-\frac{1}{2}} &= 0 \text{ km} \\
r_{64\frac{1}{2}} &= 64 \text{ km}
\end{align*}
\]

\[
\begin{align*}
\Delta z_k &= \Delta z = 1.38 \text{ km} \\
z_k &= (k - \frac{1}{2}) \Delta z \\
z_{k+\frac{1}{2}} &= k \Delta z \\
z_{-\frac{1}{2}} &= 0 \text{ km} \\
z_{65\frac{1}{2}} &= 65 \text{ km}
\end{align*}
\]

The cell boundaries at \( r_{-\frac{1}{2}}, r_{64\frac{1}{2}}, z_{-\frac{1}{2}}, z_{65\frac{1}{2}} \) coincide with the overall system boundaries.

In order to make the problem discrete in the time domain, we define a variable time step \( \Delta t_n \), so that the discrete simulation time indexed by \( n \) is given by:

\[
t_n = t_0 + \sum_{i=0}^{n-1} \Delta t_i
\]

(3.4)

where \( t_0 \) is the initial time of the simulation, generally taken to be \(-t_s\), where \( t_s \) is the duration of the cloud electrification period. A time of zero is always taken as the onset of the lightning discharge. Because the successive lightning processes of thundercloud electrification, lightning discharge, and system relaxation occur on vastly different time scales, we must make use of the variability of the time step in (3.4). Cloud electrification occurs over several seconds to several minutes, while the discharge takes place in milliseconds, and the relaxation takes many seconds. Our method to obtain stable simulation progress over this wide range of rates of change is to choose an appropriate, fixed time step in each of the three periods: electrification, discharge, and relaxation.

In this manner, the forward march of \( t_s \) versus \( n \) in (3.4) is made piecewise linear. Equivalently, we choose an appropriate number of uniform time steps in each period. Examples of two effective, discrete time formulations are shown in Table 3.1. The actual means of selecting time step length is dictated by stability considerations addressed in Section 3.2.

<table>
<thead>
<tr>
<th>Period</th>
<th>Case 1</th>
<th>Case 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Duration</td>
<td>Time steps</td>
</tr>
<tr>
<td>electrification</td>
<td>5 s</td>
<td>300</td>
</tr>
<tr>
<td>discharge</td>
<td>5 ms</td>
<td>300</td>
</tr>
<tr>
<td>relaxation</td>
<td>1 s</td>
<td>300</td>
</tr>
</tbody>
</table>

Table 3.1: Representative duration and number of time steps for the three simulation periods. Case 1 and Case 2 both involve positive cloud-to-ground discharges that lower 200 C of charge to ground.
The representative dependent variable $X(r, z, t)$ is thus made discrete in both time and space according to the following prescription:

$$X_{i,k}^n = X(r_i, z_k, t_n)$$

(3.5)

A comprehensive list of the dependent variables employed in the simulation is presented in Table 3.2. The variables are classified as either state, flux, or intermediate. State variables are those densities which the finite differences transport routine has the principal duty of advancing from one time step to the next. The flux variables are those velocities and current densities that are directly employed by the finite differences transport scheme in advancing the state variables. Finally, the intermediate variables comprise the many quantities that must be calculated at each point in time to determine the flux variables.

<table>
<thead>
<tr>
<th>Class</th>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>state</td>
<td>$n_p, n_m, n_e$</td>
<td>number density of small positive ions, small negative ions, cluster ions</td>
</tr>
<tr>
<td></td>
<td>$\rho$</td>
<td>charge density</td>
</tr>
<tr>
<td>flux</td>
<td>$\vec{u}_p, \vec{u}_m$</td>
<td>velocity of small positive ions, small negative ions</td>
</tr>
<tr>
<td></td>
<td>$\vec{J}$</td>
<td>current density</td>
</tr>
<tr>
<td>intermediate</td>
<td>$n_e$</td>
<td>number density of electrons</td>
</tr>
<tr>
<td></td>
<td>$\vec{E}_p, \vec{E}$</td>
<td>electric field, preliminary and final values</td>
</tr>
<tr>
<td></td>
<td>$T_e$</td>
<td>electron temperature</td>
</tr>
<tr>
<td></td>
<td>$T_i$</td>
<td>ion temperature (all ion species)</td>
</tr>
<tr>
<td></td>
<td>$G$</td>
<td>impact ionization rate</td>
</tr>
<tr>
<td></td>
<td>$\frac{\partial \rho_s}{\partial t}$</td>
<td>cloud charge source rate</td>
</tr>
<tr>
<td></td>
<td>$S_p, S_m$</td>
<td>all-source production rates for small positive ions, and small negative ions</td>
</tr>
<tr>
<td></td>
<td>$\tau_e, \tau_p, \tau_m$</td>
<td>mean time between collisions for electrons, small positive ions, small negative ions</td>
</tr>
<tr>
<td></td>
<td>$\mu_e$</td>
<td>electron mobility</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>total conductivity of medium</td>
</tr>
</tbody>
</table>

Table 3.2: Definitions of simulation variables.

Given that the dependent variables listed in Table 3.2 are all that are of interest, we outline the manner in which they are determined and advanced from one time step to the next. We begin with known values of the state variables. These may have been determined from the initial conditions; namely the fair weather result (2.27) for the charge density and the chemical
equilibrium result (2.117) for the particle densities. Alternatively, the state variables may have been determined by advancement from the previous time step.

Given the values of the state variables at time step \( n \), our first act is to determine the electron density from the definition of charge density (2.126):

\[
\left\{ n_{p,i,k}^n, n_{e,i,k}^n, n_{m,i,k}^n, \rho_{i,k}^n \right\} \Rightarrow n_{e,i,k}^n
\]  

(3.6)

Our next step is to use the present values of charge density to compute a preliminary solution for the electric field using Poisson Equation (2.20) as discussed in Section 2.4. This solution will be refined at high altitudes prior to computing the flux variables.

\[
\rho_{i,k}^n \Rightarrow \vec{E}_{p,i,k}^n
\]  

(3.7)

This preliminary electric field is then used to compute electron temperature throughout the simulation volume using the steady state electron energy transport equation (2.164g), all the parameters of which are specified self-consistently in terms of the electron temperature. The inelastic loss term and impact ionization appearing in (2.164g) are calculated from the collisional effects model of Section 2.5, which is itself parameterized by the background neutral densities, modeled in Section 2.2.1. Solving (2.164g) is an exercise in nonlinear root finding [Press et al., 1994, p. 340].

\[
\left\{ n_{e,i,k}^n, \vec{E}_{p,i,k}^n \right\} \Rightarrow T_{e,i,k}^n
\]  

(3.8)

Having now found the electron temperature, and knowing that the ion temperature is fixed, equal to that of the neutral background (see discussion of [2.162]), we use the collisional effects model of Section 2.5 to compute the mean collision times of all the species. This calculation, summarized in (2.97) also uses the local neutral densities of Section 2.2.1.

\[
\left\{ T_{e,i,k}^n, T_{i,k}^n \right\} \Rightarrow \left\{ \tau_{e,i,k}^n, \tau_{p,i,k}^n, \tau_{m,i,k}^n \right\}
\]  

(3.9)

It is now a straightforward matter to compute the electron mobility and the total conductivity of the medium using (2.139) and (2.140).

\[
\left\{ n_{e,i,k}^n, n_{p,i,k}^n, n_{m,i,k}^n, \tau_{e,i,k}^n, \tau_{p,i,k}^n, \tau_{m,i,k}^n \right\} \Rightarrow \left\{ \mu_{e,i,k}^n, \sigma_{i,k}^n \right\}
\]  

(3.10)

This conductivity is of particular interest at high altitude, where it is used in the high altitude electric field solution discussed in Section 2.4.3. The preliminary, Poisson solution for the electric field is used, together with the conductivity to compute the final electric field solution.

\[
\left\{ \sigma_{i,k}^n, \vec{E}_{p,i,k}^n \right\} \Rightarrow \vec{E}_{e,i,k}^n
\]  

(3.11)
Having found the electric field throughout the simulation volume, we are in position to compute the ion velocities in the present time step by evolving them from the velocities of the previous time step using a separate and subsidiary finite differences transport scheme. As discussed in Section 3.2.5, we will use the Flux Corrected Transport method for this calculation, based upon ion momentum equations in conservation law form (2.131).

\[
\begin{align*}
\{ n_{p,i,k}^n, \bar{u}_{p,i,k}^n, E_{i,k}^n \} & \Rightarrow \bar{u}_{p,i,k}^n \\
\{ n_{m,i,k}^n, \bar{u}_{m,i,k}^n, E_{i,k}^n \} & \Rightarrow \bar{u}_{m,i,k}^n
\end{align*}
\] (3.12)

We now make use of (2.139) and (2.140) to write current density in the form:

\[
\vec{J} = en_p \bar{u}_p - en_m \bar{u}_m + en_e \mu_e |\vec{E}|
\] (3.13)

The final intermediate variables required before we can advance the state variables are the source terms in the continuity equations of (2.164). These include the sources derived from the chemical effects model of Section 2.6.2 and the cloud source model of Section 2.3. Moreover, we showed in Section 2.7.1 that the evolution of the cluster ion density is governed entirely by the chemical effects model. Thus, as we integrate the rate equations (2.110) of the chemical effects model to find \( S_\sigma \) (see Section 3.4), we simultaneously advance the solution for \( n_{e} \) to the next time step.

\[
\begin{align*}
\{ n_{e,i,k}^n, n_{p,i,k}^n, n_{m,i,k}^n, n_{c,i,k}^n, T_{e,i,k}^n, T_{i,k}^n \} & \Rightarrow \{ S_{p,i,k}^n, S_{m,i,k}^n, n_{c,i,k}^{n+1} \} \\
\text{cloud source model} & \Rightarrow \frac{\partial S_{\sigma}}{\partial t}{|_{i,k}}
\end{align*}
\] (3.14)

(3.15)

The FCT algorithm discussed in Section 3.2 is now used to separately advance each of the remaining state variables, \( \rho, n_p, \) and \( n_m, \) to the next time step.

\[
\begin{align*}
\{ n_{p,i,k}^n, \bar{u}_{p,i,k}^n, S_{p,i,k}^n, \frac{\partial S_{\sigma}}{\partial t}{|_{i,k}} \} & \Rightarrow n_{p,i,k}^{n+1} \\
\{ n_{m,i,k}^n, \bar{u}_{m,i,k}^n, S_{m,i,k}^n \} & \Rightarrow n_{m,i,k}^{n+1} \\
\{ \rho_{i,k}, J_{i,k}^n \} & \Rightarrow \rho_{i,k}^{n+1}
\end{align*}
\] (3.16)

(3.17)

(3.18)

The values of the state variables are now known at time step \( n + 1 \). We use these results to start the next iteration of the solution process, beginning at (3.6) once again. Subsequent iterations advance the solution through the time interval required.
3.2 The Flux-Corrected Transport Method

Our principal means of calculating the evolution of the particle densities in the simulation is a finite differences scheme known as the flux-corrected transport (FCT) method. This approach was developed by Boris and Book [1973] and refined by Zalesak [1979]. It is one of a vast catalog of finite difference techniques. The decision to apply it to the problem at hand was based upon the fact that it is especially suited to situations involving particle flows and shocks in the fluid medium [Gatsonis, 1991]. We anticipate that the flow of particles in response to lightning-induced electric fields will play a key role in the physical effects of lightning on the mesosphere and lower ionosphere. Moreover, any system governed by conservation laws, such as our particle continuity equations (2.121), is liable to produce shocks and solitary waves [Tidman and Krall, 1971]. Indeed, it will turn out that the simulation does predict shock-like, propagating structures in the particle densities (see Chapter 4)—validating our selection of the FCT method.

We will begin this section by presenting an overview of the FCT method, with particular attention to the mechanisms by which it is well-suited to the study of particle flows and shocks. We then will examine the separate low-order and high-order finite difference schemes that are used as part of the FCT technique. We extend the method for application to cylindrical systems and momentum equations. We conclude by examining the stability and accuracy of the method.

3.2.1 Overview of FCT Method

The essence of the flux-corrected transport method is that it linearly combines a low-order and a high-order finite difference algorithm for the purpose of reproducing as accurately as possible any discontinuities that may be present in the exact solution, but without introducing ripples in the vicinity of those discontinuities.

The order of a finite difference scheme refers to the order of its truncation error, where the truncation error is the difference between the finite differences representation and the partial differential equation it represents [Anderson et al., 1984, p.46]. This order is expressed as a power of the grid spacing. Thus, the truncation error falls off more rapidly with decreasing grid spacing for a high-order scheme than for a low-order scheme. We are thereby drawn to higher-order schemes when we seek to more accurately reproduce features in the exact solution for a fixed grid spacing.
Unfortunately, high-order finite difference methods—second order methods in particular—introduce artificial, numerical dispersion that tends to produce ripples near any discontinuities in the exact solution [Anderson et al., 1984, p.92]. First-order methods, on the other hand, introduce numerical viscosity, which tends to reduce any gradients in the solution, regardless of whether their origin is physical or numerical. In fact, a first-order solution is guaranteed to be monotonic [Anderson et al., 1984, p.144]. In choosing between first- and second-order finite difference methods, we are thus faced with a tradeoff of sharpness against smoothness in any modeled transition.

It is the premise of the FCT approach that we can combine a low- and a high-order method in such a way that the smoothing effect of the low-order method counteracts the rippling effect of the high-order method, and the end result is the ability to reproduce discontinuities that are simultaneously sharp and ripple-free. This is an important characteristic in a method intended for applications that may involve shocks.

We can use a one-dimensional example to illustrate the general concept of combining two finite difference algorithms, as well as illustrate the specific FCT approach. A generic, one-dimensional continuity equation appears:

\[
\frac{\partial w}{\partial t} + \frac{\partial f}{\partial x} = 0
\]

(3.19)

where \(w\) is the density variable, and \(f\) is the flux. We represent this in finite difference form as:

\[
\frac{w_i^{n+1} - w_i^n}{\Delta t_n} + \frac{F_{i+\frac{1}{2}}^n - F_{i-\frac{1}{2}}^n}{\Delta x_i} = 0
\]

(3.20)

where the fluxes at the cell boundaries, \(F_{i+\frac{1}{2}}\), are some function of several of the grid point fluxes, \(f_i^n\), not necessarily even from the same time step, \(n\). The manner in which the \(F_{i+\frac{1}{2}}\) are calculated from the \(f_i^n\) determines the order of the scheme. One example of a first order method would be simply

\[
F_{i+\frac{1}{2}}^{LO} = f_i^n
\]

(3.21)

Based upon the Taylor series expansion of \(f\) on a uniform grid [Abramowitz and Stegun, 1972]:

\[
f(x + \Delta x) = f(x) + \Delta x \frac{df}{dx} + \frac{(\Delta x)^2}{2} \frac{d^2 f}{dx^2} + \frac{(\Delta x)^3}{6} \frac{d^3 f}{dx^3} + \cdots
\]

(3.22)

we can write the finite differences approximation to the divergence as:
\[ \frac{d f}{dx} \equiv \frac{F_{i+\frac{1}{2}}^{LO} - F_{i-\frac{1}{2}}^{LO}}{\Delta x} = \frac{f_i - f_{i-1}}{\Delta x} = \frac{d f}{dx} - \frac{\Delta x}{2} \frac{d^2 f}{dx^2} + \ldots \] (3.23)

The truncation error involved in using (3.23) is first-order in \( \Delta x \). By the same token, a simple example of a second-order scheme would have:

\[ F_{i+\frac{3}{2}}^{HI} = \frac{1}{2} \left( f_i^n + f_{i+1}^n \right) \] (3.24)

since then:

\[ \frac{d f}{dx} \equiv \frac{F_{i+\frac{3}{2}}^{HI} - F_{i-\frac{1}{2}}^{HI}}{\Delta x} = \frac{f_{i+1} - f_{i-1}}{2\Delta x} = \frac{d f}{dx} + \frac{(\Delta x)^2}{6} \frac{d^3 f}{dx^3} + \ldots \] (3.25)

which is second-order in \( \Delta x \). More sophisticated approaches, which depart from (3.21) and (3.24) in order to achieve some special suitability to a particular problem, nevertheless exhibit truncation errors of the form suggested by (3.23) and (3.25).

Combining a low-order with a high-order method thus amounts to forming a linear combination of two different expressions for the cell boundary flux:

\[ F_{i+\frac{1}{2}} = \left( 1 - C_{i+\frac{1}{2}} \right) F_{i+\frac{1}{2}}^{LO} + C_{i+\frac{1}{2}} F_{i+\frac{3}{2}}^{HI} \] (3.26)

The FCT method is defined by the manner in which it creates this linear combination. We are now in a position to sketch the actual procedure:

(a) Compute the low- and high-order cell boundary fluxes, \( F_{i+\frac{1}{2}}^{LO} \) and \( F_{i+\frac{3}{2}}^{HI} \), respectively, using any appropriate low- and high-order methods. The low-order method must be guaranteed to be dissipative and provide only monotonic solutions.

(b) Advance the solution for \( w \) using (3.20) and the low order flux:

\[ \frac{f_{i+\frac{1}{2}}^{TD}}{\Delta x_i} = f_i^n - \frac{\Delta f_i}{\Delta x_i} \left( F_{i+\frac{1}{2}}^{LO} - F_{i-\frac{1}{2}}^{LO} \right) \] (3.27)

and thereby obtain the so-called “transported and diffused” solution.

(c) Define an “anti-diffusive” flux, \( A_{i+\frac{1}{2}} \), according to:

\[ A_{i+\frac{1}{2}} = F_{i+\frac{3}{2}}^{HI} - F_{i+\frac{1}{2}}^{LO} \] (3.28)

with which to combat the diffusive effects of the low-order solution.

(d) Correct, or limit, the anti-diffusive flux using correction factors \( C \) determined according to the discussion below.

\[ A_{i+\frac{1}{2}}^C = C_{i+\frac{1}{2}} A_{i+\frac{1}{2}} \text{ where } 0 \leq C_{i+\frac{1}{2}} \leq 1 \] (3.29)
The name of the FCT algorithm derives from this "flux correction" step.

(e) Finally, apply the corrected, anti-diffusive flux in advancing the transported and diffused solution from Step (b).

\[ w_{i+1}^{TD} = w_i^{TD} - \frac{\Delta t_n}{\Delta x_i} \left( A_{i+\frac{1}{2}y_i} - A_{i-\frac{1}{2}y_i} \right) \]  

(3.30)

This is the desired, time-advanced solution for \( w \).

The five-step FCT procedure outlined above can be generalized immediately for use in two- and three-dimensional problems, simply be adding in the fluxes across the new cell boundaries that appear in higher dimensions. However, the generalization of the flux correction step (d) is nontrivial. The original algorithm for flux-limiting in one-dimension was obtained by Boris and Book [1973]. The generalization to two dimensions was accomplished by Zalesak [1979], and to three dimensions by Gatsonis [1991]. Our interest is in the two-dimensional version, generalized slightly from rectangular to cylindrical coordinates, as discussed in Section 3.2.2.

The purpose of the flux correction step is to ensure that the application of the anti-diffusive flux in Step (e) above will not introduce new extrema into the solution for \( w \). To this end, we identify a value of \( w_i \) that we wish not to overshoot and another that we wish not to undershoot, \( w_i^{max} \) and \( w_i^{min} \), respectively. It should then be true that

\[ w_i^{min} \leq w_i^{TD} \leq w_i^{max} \]  

(3.31)

In the case of the Boris and Book [1973] flux limiter, these values are obtained according to:

\[ w_i^{min} = \min\{w_{i-1}^{TD}, w_{i}^{TD}, w_{i+1}^{TD}\} \quad \text{and} \quad w_i^{max} = \max\{w_{i-1}^{TD}, w_{i}^{TD}, w_{i+1}^{TD}\} \]  

(3.32)

According to Zalesak [1979], we use \( w_i^{max} \) and \( w_i^{min} \) to compute six additional quantities at each grid point, from which we derive the correction factors, \( C \). These quantities are listed below, along with their interpretation:

\[ P_i^* = \max\left(0, A_{i-\frac{1}{2}y} \right) - \min\left(0, A_{i+\frac{1}{2}y} \right) \geq 0 \quad \text{sum of all anti-diffusive fluxes directed into cell} \ \ i \]

\[ Q_i^* = \frac{\Delta x_i}{\Delta t_n} \left( w_i^{max} - w_i^{TD} \right) \geq 0 \quad \text{upward leeway for application of anti-diffusive flux in cell} \ \ i \]

\[ R_i^* = \min\left(1, \frac{Q_i^*}{P_i^*} \right) \]
\[ P_i^- = \max(0, A_{i+\frac{1}{2}}) - \min(0, A_{i-\frac{1}{2}}) \geq 0 \quad \text{sum of all anti-diffusive fluxes directed out of cell } i \]
\[ Q_i^- = \frac{\Delta x_i}{\Delta t_i} \left( w_i^{TD} - w_i^{min} \right) \geq 0 \quad \text{downward leeway for application of anti-diffusive flux in cell } i \]
\[ R_i^- = \min \left( 1, \frac{Q_i^-}{P_i^-} \right) \quad (3.33) \]

By its construction, if \( R_i^- \) is used to multiply all anti-diffusive fluxes directed inward to cell \( i \), then we are guaranteed not to overshoot \( w_i^{max} \). Likewise, if we multiply all anti-diffusive fluxes directed outward from cell \( i \) by the factor \( R_i^- \), then we are guaranteed not to undershoot \( w_i^{min} \). Moreover, since the anti-diffusive flux through a particular cell boundary is necessarily directed out of one cell and into the adjacent cell, it is subject to two constraints, based upon avoiding the formation of an extremum in either cell. The prescription for the correction factors \( C \) that guarantees neither \( w_i^{max} \) nor \( w_i^{min} \) is overshot is just:

\[ C_{i+\frac{1}{2}} = \begin{cases} \min(R_i^-, R_{i+1}^+) & \text{for } A_{i+\frac{1}{2}} \geq 0 \\ \min(R_i^-, R_{i+1}^-) & \text{for } A_{i+\frac{1}{2}} < 0 \end{cases} \quad (3.34) \]

This completes the description of the FCT method in a rectangular coordinate system, in the absence of particle sources and sinks. For simplicity of presentation, we have considered a one-dimensional system. However, the quantities listed in (3.33) are contrived to be entirely analogous to those used by Zalesak [1979] in a two-dimensional formulation. The essential difference is simply that in two dimensions each cell has more cell boundaries to account for.

In the presence of particle production and loss, our generic continuity equation (3.19) becomes

\[ \frac{\partial w}{\partial t} + \frac{\partial f}{\partial x} = S(x, t) \quad (3.35) \]

where \( S \) represents the number of particles produced per unit length per unit time. The finite differences approximation to this equation is written:

\[ w_i^{n+1} = w_i^{TD} - \frac{\Delta t_n}{\Delta x_i} \left( A_{i+\frac{1}{2}}^C - A_{i-\frac{1}{2}}^C \right) + S_i^n \Delta t_n \quad (3.36) \]

It is appropriate to add the particle source term to the transported and diffused solution after the flux correction step of the FCT process. This is because the flux correction method is intended to prevent extrema of numerical origin from creeping into the solution. While extrema can certainly
be produced by local particle sources, they are of physical origin and ought to be preserved in the final solution.

### 3.2.2 Extending FCT to Cylindrical Systems

We can further extend the FCT solution (3.36) to the case of cylindrically symmetrical geometries. Cylindrical symmetry allows us to describe a three-dimensional physical space using a two-dimensional coordinate grid. Our point of departure is thus the two-dimensional FCT result of Zalesak [1979] in rectangular \((x, y)\) coordinates:

\[
w_{i,j}^{n+1} = w_{i,j}^{TD} - \frac{\Delta t_n}{\Delta x_i} \left( A_{i+1/2,j}^C - A_{i-1/2,j}^C \right) - \frac{\Delta t_n}{\Delta y_j} \left( A_{i,j+1/2}^C - A_{i,j-1/2}^C \right)
\]

which approximates the exact partial differential equation:

\[
\frac{\partial w}{\partial t} + \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} = 0
\]

As noted above, (3.37) is a straightforward extension of the one-dimensional result (3.30), achieved by the addition of \(y\)-directed fluxes. The only additional comment necessary is that \(A_{i+1/2,j}^C\) and \(A_{i,j+1/2}^C\) are in fact two entirely separate arrays, representing the corrected, anti-diffusive fluxes directed through all of the cell boundaries normal to the \(x\)- and \(y\)-axes, respectively. They are based on raw anti-diffusive fluxes defined in precise analogy to the one-dimensional versions:

\[
A_{i+1/2,j} = F_{i+1/2,j}^{HI} - F_{i+1/2,j}^{LO}
\]

\[
A_{i,j+1/2} = G_{i,j+1/2}^{HI} - G_{i,j+1/2}^{LO}
\]

The analog to (3.38) in the case of axisymmetric cylindrical coordinates, and including a particle source is [Beyer, 1981, p.364]:

\[
\frac{\partial w}{\partial t} + \frac{\partial f}{\partial r} + \frac{\partial g}{\partial z} + \frac{f}{r} = S(r, z, t)
\]

Since the additional term that arises in the case of cylindrical geometry, the geometric divergence, is known exactly at the center of each grid cell, the appropriate finite differences approximation to this equation is just:

\[
w_{i,k}^{n+1} = w_{i,k}^{TD} - \frac{\Delta t_n}{\Delta r_i} \left( A_{i+1/2,k}^C - A_{i-1/2,k}^C \right) - \frac{\Delta t_n}{\Delta z_k} \left( A_{i,k+1/2}^C - A_{i,k-1/2}^C \right) - \frac{f_{i,k}^n}{r_i} \Delta t_n + S_{i,k}^n \Delta t_n
\]
Here $f_{i\ell}^n$ is the raw, radial flux density at the center of each grid cell. The anti-diffusive fluxes are defined according to (3.39), and the flux correction is accomplished according to Zalesak [1979], in a procedure similar to that outlined in (3.33). The flux corrected transport algorithm employed in the present simulation is that of (3.41).

3.2.3 Low-order Algorithm

In order to complete the specification of our FCT approach, we must also specify the actual algorithms used to calculate the low- and high-order fluxes appearing in (3.39). There are a great many algorithms to choose from, and the selection of a particular method is based on physical considerations. We follow the work of Gatsonis [1991] in this regard.

The low-order method used here is known as the donor cell algorithm. The key feature of this algorithm is the use of the upwind density in computing the flux at each grid cell interface. This means that the density at an interface is taken to be the density from the previous time step on the upwind side of the interface. The donor cell approach to specifying the particle density at the interface is plainly motivated by physical consideration of particle flow. If $u$ represents particle velocity in the $r$-direction, then:

$$w_{i\pm\frac{1}{2},k}^{DC} = \begin{cases} w_{i,k}^{n-1} & \text{if } u_{i+\frac{1}{2},k}^n \geq 0 \\ w_{i+1,k}^{n-1} & \text{if } u_{i+\frac{1}{2},k}^n < 0 \end{cases}$$  \hspace{1cm} (3.42)

where the velocity at the interface is just the average velocity from the two adjacent cells:

$$u_{i\pm\frac{1}{2},k}^n = \frac{1}{2} \left( u_{i,k}^n + u_{i+1,k}^n \right)$$  \hspace{1cm} (3.43)

Similar results hold for the interfaces normal to the $z$-axis. The low-order fluxes may then be written

$$F_{i+\frac{1}{2},k}^{LO} = w_{i+\frac{1}{2},k}^{DC} u_{i+\frac{1}{2},k}^n$$  \hspace{1cm} (3.44)

$$G_{i,k+\frac{1}{2}}^{LO} = w_{i,k+\frac{1}{2}}^{DC} v_{i,k+\frac{1}{2}}^n$$  \hspace{1cm} (3.45)

We recognize from (3.21) that (3.42) is essentially a first order specification of the interface density. On the other hand, by comparison with (3.24) we see that the interface velocity (3.43) is second-order. Accordingly, Gatsonis [1991] reports that the accuracy of the method is more than first order, but it is nevertheless dissipative and suitable for use as the low-order algorithm in the FCT scheme.
Moreover, Gatsonis [1991] performs Von Neumann stability analysis on the method. Von Neumann analysis is a standard means of examining the amplification of round-off error as a finite differences solution marches from one time step to the next under the influence of constant convective velocities [Anderson et al., 1984, p.71]. Gatsonis [1991] concludes that stability is ensured if

$$c_x + c_y \leq 1$$  \hspace{1cm} (3.46)

where $c_x$ and $c_y$ are the Courant numbers defined by:

$$c_x = \frac{|u|\Delta t}{\Delta x} \quad \text{and} \quad c_y = \frac{|v|\Delta t}{\Delta y}$$  \hspace{1cm} (3.47)

and $u$ and $v$ are the x- and y-velocities in the rectangular system. Rewriting this for our cylindrical system, we find that the condition for global stability of the low-order solution is

$$\max_{i,k} \left\{ \left| u_{r,i,k}^{n} \right| \frac{\Delta t_n}{\Delta r_i} + \left| u_{z,i,k}^{n} \right| \frac{\Delta t_n}{\Delta z_k} \right\} \leq 1$$  \hspace{1cm} (3.48)

This ensures that (3.46) is not violated anywhere in the simulation volume. (3.48) provides a condition on temporal step size and grid spacing such that stability is maintained. In particular, we see that for a given temporal step size, grid spacing must be larger than some minimum to ensure stability. By the same token, if grid spacing is fixed, then the temporal step size must be smaller than some maximum value to ensure stability. In practice, we have determined the grid spacing rather solidly and will use (3.48) to govern the time steps.

### 3.2.4 High-order Algorithm

The high-order scheme proposed by Zalesak [1979] is of the predictor-corrector type. The leapfrog method serves as the predictor, while the trapezoidal serves as the corrector. The predictor step provides the densities at provisional time step $(n+1)$, by advancing the solution from the previous time step $(n-1)$:

$$w_{i,k}^{n+1} = w_{i,k}^{n} - \frac{2\Delta t_n}{\Delta r_i} \left( F_{i+\gamma,k}^{n} - F_{i-\gamma,k}^{n} \right) - \frac{2\Delta t_n}{\Delta z_k} \left( G_{i,k+\gamma}^{n} - G_{i,k-\gamma}^{n} \right)$$  \hspace{1cm} (3.49)

where the fluxes $F$ and $G$ are obtained by a fourth order method from the raw fluxes [Gatsonis, 1991, p.98]. These densities are used to calculate the raw fluxes at the provisional time step, $f^P$ and $g^P$, which are then used in the trapezoidal corrector step. We first form:
\[ f_{i,k}^* = \frac{1}{2} \left( f_{i,k}^n + f_{i,k}^p \right) \quad \text{and} \quad g_{i,k}^* = \frac{1}{2} \left( g_{i,k}^n + g_{i,k}^p \right) \quad (3.50) \]

We use these to calculate the associated fourth-order flux approximations, and then advance the solution using the corrector step:

\[ w_{i,k}^{n+1} = w_{i,k}^n - \frac{\Delta t_n}{\Delta r_i} \left( F_{i+\frac{1}{2},k}^* - F_{i-\frac{1}{2},k}^* \right) - \frac{\Delta t_n}{\Delta z_k} \left( G_{i+\frac{1}{2},k}^* - G_{i-\frac{1}{2},k}^* \right) \quad (3.51) \]

The stability criterion obtained by Von Neumann analysis for the predictor step is:

\[ \max_{i,k} \left( \left| u_{r,i,k}^n \right| \frac{\Delta t_n}{\Delta r_i} + \left| u_{z,i,k}^n \right| \frac{\Delta t_n}{\Delta z_k} \right) \leq 1 \quad (3.52) \]

the same as for the low-order method. The same analysis applied to the corrector step yields a slightly more relaxed criterion:

\[ \max_{i,k} \left( \left| u_{r,i,k}^n \right| \frac{\Delta t_n}{\Delta r_i} + \left| u_{z,i,k}^n \right| \frac{\Delta t_n}{\Delta z_k} \right) \leq \sqrt{2} \quad (3.53) \]

We apply the more restrictive condition (3.52) to ensure the stability of the overall FCT method.

### 3.2.5 Application of the FCT Method to the Momentum Equation

The flux-corrected transport method outlined in Section 3.2.1 can be applied to any hyperbolic equation in conservation law form; that is, in the form of (3.35). As such, it may be applied directly to each component of the momentum equation (2.131), reproduced below with the right-hand side combined into a single source vector.

\[ \frac{\partial}{\partial t} (n_a \bar{u}_a) + \bar{V} \cdot (n_a \bar{u}_a \bar{u}_a) = \bar{S}_a \quad (3.54) \]

If we define:

\[ w = n_a u_{az}; \quad f = w u_{ar}; \quad \text{and} \quad g = w u_{az} \quad (3.55) \]

then we can express the \( z \)-component of (3.54) as:

\[ \frac{\partial w}{\partial t} + \frac{\partial f}{\partial r} + \frac{\partial g}{\partial z} + \frac{f}{r} = S_{az} \quad (3.56) \]

and similarly for the \( r \)-component. We note that (3.56) is now in the familiar, conservation law form of (3.40), and the FCT method may be applied precisely as before.
3.2.6 Stability and Accuracy of the FCT Method

The importance of the stability criteria (3.48), (3.52) and (3.53) stems from two sources. First, as a practical matter, stepping the solution under circumstances that violate these criteria quickly leads to exploding particle densities and numerical overflows, as rounding errors amplify out of control. Second, stability is one of the two requirements for an effective finite differences scheme [Anderson et al., 1984, p.46]. The other is consistency, the property that the finite differences approximation approaches the original partial differential equation as the spatial and temporal mesh spacing is reduced to zero. This property is satisfied when the truncation error is composed of terms that are proportional to some power of the time step size or spatial grid spacing. In fact, consistency is a feature of most rationally conceived finite difference schemes, including the donor cell, leapfrog, and trapezoidal algorithms. This leaves stability as the primary concern.

As noted in Section 3.1, the spatial grid spacing was fixed based upon the physical scale lengths of the phenomena and ambient environment under study. In order to maintain the stability of the FCT method, the time step size is adjusted to be just small enough to satisfy the most restrictive of the stability criteria, (3.52), for the chosen grid spacing. The operation of the code is such that the simulation step size is adjusted adaptively so as to keep the left-hand side of (3.52) in the range between 0.2 and 0.8. Once we implement a control such as this, stability is assured, and we turn our attention to the accuracy of the method.

Any examination of the accuracy of the FCT method must bear in mind that the local accuracy of the time-advanced solution is less important than the accuracy of the total particle content of the system. Consider, for example, Figure 3.1a. This figure shows a time series of cuts along the simulation axis of particle density in a cylindrical blob of particles that is propagating uniformly, upward and out of the simulation volume at a speed of 0.78 km/s. The profile of the solution becomes increasingly rounded with continued propagation. This rounding is a purely numerical effect. By the time the blob has propagated for 30 s, its amplitude is reduced to 97% of its original value. If we were to compare this FCT result with the corresponding result obtained by using the low-order donor cell algorithm alone, we would find significantly more rounding with the donor cell method, which yields a 93% amplitude after 30 s of propagation. Local errors on the order of several percent are not unreasonable given the fact that Figure 3.1 represents an extreme
Figure 3.1: Test run of cylindrical particle blob propagating in uniform velocity field using flux-corrected transport (FCT) algorithm. Radius of blob is 5 km, initial altitude extent is 15 km to 30 km. Propagation velocity is vertically upward at a velocity such that the Courant number is 0.51. (a) Altitude profiles of density along simulation axis as a function of increasing time. (b) Total number of particles in cylindrical simulation volume. This quantity is constant to within machine rounding error until the blob begins to exit the simulation volume.
case of extended propagation over many 10’s of seconds. However, if the 3% decline in density were to stem from a loss of particles, then the consequences to the simulation would be severe.

In particular, if our numerical method did not conserve the total amount of charge when advancing the charge density, then total charge would be subject to numerical inaccuracies, leading to erroneous electric fields at the extremities of the system. The consequences of this type of error are much more severe than those of minor pulse reshaping, the effects of which are purely local. Thus, the question of whether or not the FCT method is conservative is an important one.

Applying Gauss’s Law to (3.40), we obtain:

\[
\frac{\partial}{\partial t} \int_V w \, d^3r + \oint_{\partial V} \mathbf{\hat{n}} \cdot \mathbf{\hat{I}} \, dA = \int_S d^3r
\]  

(3.57)

where \( V \) is some volume in the system, \( \partial V \) is the surface that bounds it, \( \mathbf{\hat{n}} \) is a unit vector normal to \( \partial V \), \( dA \) is an area element, and \( \mathbf{\hat{I}} = f \mathbf{\hat{r}} + g \mathbf{\hat{z}} \). According to (3.57), in the absence of a production source, and flux at the boundary, the integral of the density \( w \) is constant.

\[
\int_V w \, d^3r = \text{constant}
\]  

(3.58)

In the example of Figure 3.1, the property (3.58) should be maintained as long as the particle blob stays away from the boundary of the integration volume \( V \). We verify that this is the case by performing a two-dimensional numerical integration of the predicted particle density over the entire simulation volume at each time step. The method of integration is chosen to make use only of the tabulated simulation values at the grid points—instead of interpolated values at arbitrary locations in \( (r, z) \). This limits our choice of integration methods to the most elementary ones, and any of these will serve equally well. We use the technique of repeated one-dimensional integration, based upon a closed, extended formula for a uniform grid given by Press et al. [1994, p. 128]:

\[
\int_{x_1}^{x_2} f(x) \, dx = \Delta x \left[ \frac{1}{6} f_1 + \frac{7}{6} f_2 + \frac{11}{24} f_3 + f_4 + f_5 + \cdots + f_{N-3} + \frac{11}{24} f_{N-2} + \frac{7}{6} f_{N-1} + f_N \right]
\]  

(3.59)

where \( f_n = f(x_n) \). The error term associated with (3.59) is of order \( N^4 \), but again we are not concerned with the absolute accuracy of the calculated integral, only whether it remains constant.

Figure 3.1b shows the time series of the integration results and demonstrates that the total particle content of the system is conserved by the FCT method. In fact, the total particle content is maintained to within the rounding error of the computer. The decrease shown for late times is due
to the particle blob exiting the system. For all practical purposes, then, the FCT method maintains
the total particle content of the system as an exact constant.

3.3 Finite Differences Solution of the Poisson Equation in Cylindrical Coordinates

At each time step taken by the flux-corrected transport algorithm, the electric field vector
must be computed at every grid point in the simulation volume. An effective numerical method for
the solution of Poisson's Equation in a cylindrical system is therefore a critical component of the
simulation. The method we will employ was developed by Swarztrauber and Sweet [1975] as part
of the FISHPAK package of FORTRAN routines for the solution of elliptic partial differential
equations. We will briefly summarize the method, and then examine its accuracy.

3.3.1 FISHPAK Subroutine PWSCYL

The FISHPAK subroutine PWSCYL [Swarztrauber and Sweet, 1975] solves a finite
differences approximation to the Helmholtz equation in cylindrical coordinates. It accommodates
uniform spatial grids of user-specified size. It treats a large number of possible boundary
conditions. In short, it offers sufficient flexibility for us to use it without any modification.

The spatial grid is the same as that used by the transport solution, so that charge density
can be transferred to PWSCYL without the need for interpolation. The domain of the Poisson
solution is thus divided into 64 panels in the vertical dimension, and 63 panels in the radial
dimension in order to span the 65 and 64 grid points, respectively, of the transport solution.

As discussed in Section 2.4.2, we are taking the top and bottom boundaries of the
simulation volume to be perfectly conducting. Conditions at the cylindrical surface correspond to
the quiescent, fair-weather electric field. In terms of the boundary conditions applied in PWSCYL,
these translate into:

\[ \Phi(r, z = 0) = 0 \]  
(3.60)

\[ \Phi(r, z = z_{\text{max}}) = V_0 \]  
(3.61)

\[ \frac{\partial}{\partial r} \Phi(r = a, z) = 0 \]  
(3.62)

where \( \Phi \) is the solution provided by the PWSCYL subroutine and \( r = a \) describes the cylindrical
surface of the simulation volume. Thus the top and bottom of the solution domain are specified as
equipotential surfaces. The one at ground level is arbitrarily assigned an electrostatic potential of zero, while the one at ionospheric altitude is held fixed at the earth-ionosphere potential of (2.34).

The partial differential equation addressed by PWSCYL is the Helmholtz equation in axisymmetric cylindrical coordinates:

$$\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial \Phi}{\partial r} \right) + \frac{\partial^2 \Phi}{\partial z^2} + \frac{\lambda}{r^2} \Phi = P(r, z)$$  \hspace{1cm} (3.63)

In our case, the constant $\lambda = 0$, and (3.63) expresses the cylindrical version of the Poisson Equation (2.37) for an electrostatic field associated with a continuous distribution of charge in a vacuum:

$$\nabla^2 \Phi = -\frac{\rho(r, z)}{\varepsilon_0}$$  \hspace{1cm} (3.64)

We use the vacuum value of the permittivity here because all the charge in the medium is free and accounted for in $\rho$, so that there is no polarization charge in the medium. The finite differences approximation to (3.63) is obtained by inserting a centered approximation for each of the derivatives:

$$\frac{1}{(\Delta r)^2 r_i} \left[ (r_i + \frac{1}{2} \Delta r)(\Phi_{i, k+1} - \Phi_{i, k}) - (r_i - \frac{1}{2} \Delta r)(\Phi_{i, k} - \Phi_{i, k-1}) \right]$$

$$+ \frac{1}{(\Delta z)^2} \left[ \Phi_{i, k+1} - 2\Phi_{i, k} + \Phi_{i, k-1} \right] + \frac{\lambda}{r_i^2} \Phi_{i, k} = P(r_i, z_k)$$  \hspace{1cm} (3.65)

This system of equations is presumed to apply at the domain boundary $r = a$. The additional condition imposed by (3.62) is satisfied by creating virtual grid points one radial grid spacing beyond the boundary. The values of $\Phi$ at those points are then related to the rest of the unknowns by the second order difference form of (3.62).

$$\frac{\Phi_{M+1,k} - \Phi_{M-1,k}}{2 \Delta r} = 0$$  \hspace{1cm} (3.66)

where $i = M$ represents the outermost actual radial grid position. Given conditions (3.60), (3.61), and (3.62), the problem is well-posed in the case of $\lambda = 0$, and PWSCYL provides the solution $\Phi_{i, k}$ to the system of algebraic equations represented by (3.65) and (3.66).
3.3.2 Electric Field Calculation

The electrostatic potential provided by the PWSCYL solution of (3.64) is of no direct use in computing particle velocities for the transport calculation. The electric field vector is the quantity of interest at all grid points in the simulation volume. In the analytical sense, the electric field is obtained in straightforward manner from the exact electrostatic potential, \( \phi \):

\[
\vec{E} = -\nabla \phi
\]  

(3.67)

However, some sophistication is required to obtain an estimate of the electric field from our finite differences solution \( \Phi_{i,k} \).

We are concerned with estimating the electric field components at the grid points, and not between them. Moreover, we need to know the electric field even at the simulation boundary. This calls for an interpolative scheme to perform numerical differentiation on the values \( \Phi_{i,k} \) in order to estimate \( E_{r,k} \) and \( E_{z,k} \). To this end, we first consider each column of the array \( \Phi_{i,k} \) separately. We fit a polynomial of degree \( M \) to the evenly spaced values in each column and then evaluate the derivative of that polynomial at each grid point to estimate the radial component of the electric field. We perform the same operation row-by-row to obtain the vertical electric field component.

The technique of Savitzky-Golay smoothing filters provides an efficient means of computing the derivatives as a weighted sum of the grid-point values \( \Phi_{i,k} \) [Press et al., 1994, p. 644]. We apply a Savitzky-Golay filter of order \( M = 4 \) along each dimension of the electrostatic potential array to obtain the desired electric field values. The degree \( M = 4 \) was chosen after considerable experimentation with data highly representative of the \( \Phi_{i,k} \) arrays encountered in practice with the simulation.

3.3.3 Discretization Error in the Electric Field Solution

The difference between the solution of the finite difference approximation to a partial differential equation and the exact solution of the equation itself is known as the discretization error. Since these calculations are undertaken in double precision, the discretization error is much larger than any accumulated machine round-off error, and we concern ourselves only with the discretization error as the principal source of inaccuracy in our solution.

In order to characterize the discretization error, we must apply our combined Poisson solver/electric field calculation routine to a charge distribution for which the exact electric field...
solution is known. Moreover, we must accomplish this for the boundary conditions and grid geometry that we plan to use in the actual operation of the simulation. Finally, we must examine the error introduced in the particular electric field component, and at the location, that is most important to the outcome of the simulation. In practice, this means the vertical electric field component along the simulation axis.

We consider a cylindrical volume, uniformly filled with positive charge, of radius 5 km and extending upward from the ground plane to an altitude of 20 km. We plan to make use of an exact solution for the electric field due to a cylindrical volume charge in a vacuum. Thus, in order to account for the conductive planes at the ground and at 100 km altitude, we must introduce image charges. The primary image is a cylinder of negative charge extending below the ground plane. Additional images must be introduced to account for the upper conductive boundary. In fact, in the case of two parallel conductive boundaries we need an infinite series of images. The situation is illustrated in Figure 3.2a.

The elementary building block of the array of charge volumes pictured in Figure 3.2a, is the bipolar charged cylindrical volume extending 20 km above and below the ground plane. We adapt an exact solution for the vertical component of the electric field along the axis of a uniformly charged cylinder [Zahn, 1987, p. 72]. The corresponding solution for the bipolar case is:

\[
E_z^0(z) = \frac{\rho_0}{2\varepsilon_0} \left[ \left( a_0^2 + (z - H)^2 \right)^{1/2} - |z - H| + \left( a_0^2 + (z + H)^2 \right)^{1/2} - |z + H| - 2 \left( a_0^2 + z^2 \right)^{1/2} + 2z \right]
\]

(3.68)

where \( \rho_0 \) is the uniform positive charge density in the upper half of the cylinder, \( a_0 \) is the radius of the cylinder and \( H \) is the distance it extends in either direction from the \( z = 0 \) plane. We can now develop the exact solution of the presence of the conductive boundaries by superposing the fields due to an infinite number of such units, spaced as shown in Figure 3.2a. However, instead of adding an infinite number of images, we are content with the twenty nearest to the ground.

\[
E_z(z) = \sum_{i=0}^{20} E_z^0(z - h_i)
\]

(3.69)

where \( h_i \) is the altitude of each of the dashed lines in Figure 3.2a. We thereby find an approximate analytical solution to the problem. We can easily estimate the error of this approximation by checking the field contribution of the first neglected image. This image contributes a fraction
Figure 3.2: Test run of FISHPAK cylindrical Poisson solver. Boundary conditions and geometry are characteristic of simulation. (a) Construction of image charges used in obtaining exact solution. Physical charge is uniformly distributed inside a cylindrical volume of 5 km radius, between ground and 20 km altitude. Charge density is +2.6E-14 C per m³. Perfectly conducting planes are located at 0 km and 100 km altitude. A total of 20 image distributions is used. (b) Comparison of vertical electric field on simulation axis as function of altitude. Solid curve is computed by FISHPAK Poisson solver. Dashed curve is computed from exact solution. Discrepancy near cusp of exact solution is due to smoothing effect of the differentiator.
4.96E-05 to the total, which we will see is small compared to the discretization error of the main problem.

We are now in position to compare the vertical electric field from the analytical result (3.69) with the one calculated by the finite difference routine PWSCYL, followed by the Savitzky-Golay differentiator. The two results are plotted in Figure 3.2b. The only visible difference between the two solutions is near the cusp of the analytical solution, which occurs at the boundary of the charged volume. This discrepancy stems from the smoothing effect of the Savitzky-Golay filter—flattening the cusp. The differentiation scheme was not optimized for examples containing knife-edge discontinuities in charge density, as does this case. Practical examples obtained from the simulation have smooth variations in charge density and will not exhibit the cusp discrepancy shown here.

In any case, a close examination of the difference between the two solutions indicates that the magnitude of the discretization error exhibits both an additive and a multiplicative part, and may be written:

\[
\text{discretization error} = 1.4E-03 \ |E_z| + 1.7E-02 \ \text{V/m}
\]  
(3.70)

Thus, the electric field solution is in general accurate to three significant figures. Unfortunately, in regions of extremely low electric field, such as at high altitudes, the additive part of the error becomes significant. This is the reason for the dedicated high altitude electric field solution developed in Section 2.4.3.

It should be noted that (3.70) also incorporates the effects of the boundary condition (3.62), which forces the finite differences solution for the radial field to be zero at the cylindrical surface. This boundary condition is not shared by the analytical solution, nor the real-world situation in the earth ionosphere system. It is not possible to assign fractions of the error (3.70) to the numerical method and to the unphysical boundary condition. However, the generally small value of (3.70) indicates that the effect of the erroneous boundary condition on the vertical electric field is minimal. In Section 2.4.2 we concluded that the effect on the radial electric field is likewise minimal.

3.4 The Runge-Kutta Method

The requirement to integrate a system of ordinary differential equations as a function of time arises when we seek to obtain finite increments to the charged population densities from the
chemical rate equations (2.110). We make use of finite increments instead of simply inserting the rates of particle production from the right-hand sides of (2.110) into the equations of continuity (2.121). This is because the chemical effects can proceed on a time scale much shorter than that of the transport effects.

Based upon Table 2.5, the faster chemical processes, in particular three-body attachment, are associated with time scales of order 10 \( \mu \text{s} \). The transport processes affecting the system operate on time scales ranging from tens of microseconds during lightning discharge, to tens of milliseconds during electrification, according to Table 3.1. In the case of a fast chemical process at work during the electrification period, for example, the first order procedure of simply inserting the production rates as sources in the FCT continuity solver will lead to wildly erroneous results. The chemical system may start out with extremely high rates, but will approach equilibrium long before the end of the transport time step. The first order inclusion of the rates in the transport algorithm would simply project the high initial rates forward through the entire transport time step.

On the other hand, the time scales of the chemical and transport processes are in many cases comparable, so that it is equally inappropriate to solve the chemical rate equations (2.110) in the steady state. We are thus left with no choice but to integrate the system of rate equations.

Our chosen method of integration is a fifth-order, Runge-Kutta-Fehlberg formula, with adaptive step size control based upon an embedded fourth-order formula [Press et al., 1994, p. 708]. The fifth- and fourth-order results are, in effect, compared to obtain an estimate of the truncation error of the fourth-order result at each time step. Steps that fail a specified accuracy requirement are retried over a shorter time interval, down to a specified minimum interval.

Sufficient accuracy is required of each Runge-Kutta step that the accuracy of the integration over the transport time step is one part in 1.0E+09 plus 1.0E-04 per m\(^3\). This will ensure that the worst case error over 1000 transport steps is well within acceptable limits.
Chapter 4

Simulation Results

We have developed a system of models to simulate the electrostatic, transport, and chemical effects on the lower ionosphere due to a lightning discharge. We now exercise the simulation to explore both its general results and certain specific phenomena that it predicts. We first provide an overview of the outputs for a representative, positive cloud-to-ground lightning stroke, and make a comparison to the limited set of corresponding data that is available. We next examine the sensitivity of the results to changes in several basic lightning and ionospheric parameters. Finally, we present several new phenomena that are predicted by the simulation—in particular, high altitude charge separation, persistent effects due to ion relocation, and propagating features.

4.1 Overview of Representative Outputs

Our representative lightning stroke is a positive cloud-to-ground discharge that lowers 200 C of charge to ground. According to Table 1.2, this is a larger than average, but not extreme example of positive CG lightning. In any case, our goal at this point is less to examine the effects of an average lightning stroke than to examine a lightning stroke that strongly produces representative simulation results. The shape of the thundercloud charge distribution is a Gaussian monopole, with a 3 km radial scale length, and it is centered on the simulation axis, at an altitude of 15 km.

All of the simulation results presented in this chapter assume a monopolar cloud charge distribution instead of a dipolar one. The distinction between the two models is illustrated in Figure 2.5 and discussed in Section 2.3. A comparison of simulation results obtained for each of the models shows no significant difference. These results are presented in Appendix C. Consequently, we employ the monopolar cloud charge model without loss of generality.

The temporal evolution of the cloud charge begins with the electrification period. This period consists of an exponential growth in charge density throughout the cloud, lasting for 5 s,
with a 1.25 s e-folding time. The discharge initiates at \( t = 0 \) s, with a 1.25 ms e-folding time, which corresponds to a peak lightning current of 160 kA. For purposes of identifying intervals with equally spaced time steps, we end the discharge period at \( t = 5 \) ms, and begin the relaxation period, which extends to \( t = 1 \) s. The flow of lightning current is continuous through the discharge and relaxation periods, and there is no physical distinction between the two. The relaxation period is distinguished only by the fact that the rapid changes of the early discharge period have subsided, and we are able to switch to a longer time step. 300 time steps are used for each of the three periods in this example.

### 4.1.1 Spatial Variation of Observables

The maximum disturbance caused by this discharge occurs between 3 and 4 ms after discharge onset. This interval is marked by the largest values of electric field and electron temperature observed during the simulation. Figures 4.1 and 4.2 show the vertical and radial electric fields, respectively, as a function of position in the simulation volume, at 3 ms after discharge onset. The most prominent characteristic of the electric field at this time is the large field directed inward toward the position of the cloud charge center. The maximum field amplitude in all directions is roughly 70 kV/m and occurs approximately 3 km from the cloud center. The vertical electric field on the axis at an altitude of 70 km is 8 V/m.

The inward direction of the electric field indicates negative space charge at the location of the charge center. This is in spite of the fact that the cloud was the site of positive charge during the electrification and a small, but significant fraction of that charge should remain 3 ms after discharge onset. In particular, 9% of peak cloud charge, or 18 C should remain in the cloud at that time. The reason for instead finding net negative charge at that time is the response of the medium to shield the cloud charge that is introduced during electrification. During the electrification period, the electric field in the vicinity of the cloud radiates outward, drawing electrons and small negative ions downward from higher altitudes, and tending to expel ambient small positive ions. The electrons arriving in the region are quickly converted to negative ions by the process of three-body attachment. The negative charge is thereby immobilized.

We will consider the details of these effects at length in the coming sections, but the net result is that an enhanced concentration of negative ions builds up near the cloud during the electrification period. The negative ions partially neutralize the cloud charge introduced there.
Figure 4.1: Vertical electric field versus position at $t = 3$ ms after onset of positive CG lightning discharge with 200 C lowered to ground, and a peak current of 160 kA. Charge center at 15 km altitude. (a) Vertical electric field throughout simulation volume. (b) Vertical electric field along simulation axis. Broken line indicates negative field. Parallel trace indicates extent of high altitude electric field solution, beginning at 68.5 km. Maximum positive field is 7.6E+04 V/m at 12.1 km altitude. Maximum negative field is -6.5E+04 V/m at 18.8 km altitude.
Figure 4.2: Horizontal electric field versus position for conditions of Figure 4.1. (a) Horizontal electric field throughout simulation volume. Maximum field is -7.0E+04 V/m, at 14.75 km altitude, 3.75 km radius. (b) Horizontal electric field 30 km from simulation axis. Broken line indicates negative field. Parallel trace indicates extent of high altitude electric field solution, beginning at 68.5 km.
relocation of particles develops over several seconds during the electrification period. When the discharge commences, it proceeds too quickly for the reverse transport process to disperse the negative charge over a commensurate time scale. Thus, as positive charge is withdrawn from the region during the discharge period, a net negative charge remains and grows in magnitude until the chemical and transport processes combine to eliminate it.

Our steady state solution of the electron energy transport equation (2.163) relates the electric field to the electron temperature throughout the simulation volume. Figure 4.3 shows the electron temperature simultaneous with the fields of Figures 4.1 and 4.2. The maximum temperature observed at this time is 0.71 eV, at an altitude of 56.4 km on the simulation axis. At a slightly earlier time in the discharge period, this value of electron temperature was obtained at altitudes up to 72 km. The region of maximum temperatures is restricted to lower and lower altitudes with the passage of time. This is a result of the downward expansion of a region that is shielded from electric fields by the dielectric response of the medium.

The medium at the highest altitudes undergoes electrical relaxation most quickly, due to the great speed of charge transport there. Thus, lightning effects in the 70 to 90 km altitude region last milliseconds at most, but the medium’s response at lower altitudes can take tens of seconds to run its course. The temporal evolution of the system is discussed in Section 4.1.2.

As exemplified by Figure 4.3, the highest electron temperatures that result from a lightning discharge occur in the band of altitudes from 50 to 70 km. This is the inescapable result of the rapid decline of neutral density with altitude and the accompanying rise in electron mobility. The term in the electron energy transport equation (2.154) that represents the source of electron heating due to an applied electric field is:

\[-e n_e \vec{E} \cdot \vec{u}_e = e n_e \mu_e |\vec{E}|^2\]  \hspace{1cm} (4.1)

where we have used Ohm’s Law (2.139) in rewriting the electron drift velocity. Based upon this term, the strength of the heating is jointly proportional to the electron mobility and the square of the electric field magnitude. Electron mobility is an exponentially growing function of altitude throughout the simulation volume. On the other hand, the electric field produced by cloud height charge removal, and by charge separation at higher altitudes, decreases rapidly with altitude due to the shielding of the intervening plasma. The end result of these two effects is that term (4.1) will be a sharply peaked function of altitude. In other words, there is a narrow range of altitudes that
Figure 4.3: Electron temperature versus position for conditions of Figure 4.1. (a) Electron temperature throughout simulation volume. (b) Electron temperature versus altitude along simulation axis. Dashed line shows quiescent temperature versus altitude. Maximum temperature near cloud is 0.6 eV at 18.8 km altitude. Maximum high altitude temperature is 0.71 eV at 56.4 km altitude.
are high enough for the electron mobility to permit heating, but not so high that the electric field has been extinguished.

Not only does the electron temperature exhibit peaks of enhancement near 60 km and again in the region of intense electric fields just above the thundercloud, it also exhibits a general enhancement at all altitudes up to roughly 80 km, where the electric fields are utterly extinguished. This is demonstrated in Figure 4.3b, which shows temperatures at least ten times the ambient values at all altitudes up to 80 km. This general elevation of electron temperature helps to account for the decrease in conductivity that is observed at all altitudes from 25 km to 80 km, as illustrated in Figure 4.4a. Reduced conductivity in this altitude range permits the electric fields of the discharge period to penetrate to slightly higher altitudes than would otherwise be possible.

The decrease in conductivity develops during the discharge period in spite of the impact ionization that occurs during the discharge, as well as the enhancements in the negative ion density that develop during the electrification period. The products of impact ionization, along with the negative ions that accumulate during the electrostatic shielding of the electrifying cloud, all contribute to increasing the conductivity above the thundercloud. However, we see that while impact ionization does take place in the region of elevated temperature between 55 and 70 km altitude, the small number of ion pairs produced (shown in Figure 4.4b) has no perceptible effect on the total conductivity. Likewise, Figure 4.4c shows that while changes in ion density develop at all altitudes during the electrification period and persist into the discharge period, these changes are minor everywhere except in the immediate vicinity of the thundercloud. Thus, the sudden increase in electron temperature during the discharge period, coupled with the simultaneous decrease in electron density discussed below, far outweigh the effects of the modest numbers of charge carriers that are added to the region by impact ionization and the electrostatic shielding process.

According to Figure 4.4c, the strongest effects of the electrification process are felt in the small negative ion and positive cluster ion populations near and above the thundercloud. The cluster ion density is reduced in this region due to an elevated rate of recombination with electrons and small negative ions. The numbers of the negative species grow during electrification as they are transported into the region to participate in the electrostatic shielding of the cloud charge. The enhanced small negative ion density shown in Figure 4.4c is the direct consequence of this
Figure 4.4: Altitude profiles along the simulation axis of additional quantities obtained under the conditions of Figure 4.1. (a) Total conductivity. Broken line shows quiescent values; solid line shows values 3.0 ms after discharge onset. (b) Total impact ionization during one time step. Accumulated values for entire discharge period are several hundred per cubic meter. (c) Change in ion densities from quiescent values.
transport effect. At the time illustrated in the figure, the small positive ion density has largely been restored to quiescent levels by the cloud height charge removal associated with the discharge.

The large numbers of electrons and small negative ions drawn into the region between 10 and 20 km altitude during the electrification period account for the enhanced conductivity observed in this region, and demonstrated in Figure 4.4c. The combination of enhanced total conductivity at low altitudes and depressed conductivity at high altitudes persists throughout the discharge period and well into the relaxation period, due to the slow rate at which the negative species are removed by recombination.

In contrast to the restrained behavior of ions at high altitudes, electron flows produce noteworthy features above 20 km altitude during the discharge period. Significant depletions and enhancements of electron density relative to quiescent values are demonstrated in Figure 4.5. The greatest density excursion is the depletion that peaks near 60 km altitude. Based upon the equilibrium electron density shown in Figure 2.11, the depletion illustrated in Figure 4.5 amounts to roughly 2 % of the quiescent electron population. It is tempting to explain this feature as the result of electrons being swept to higher altitudes by the negative electric field that appears during the discharge period. However, the depletion cannot be explained by bulk electron transport, because electron heating impedes electron drifts in this region. In particular, the electron mobility near 60 km altitude for the temperatures observed is roughly 1.0E+03 m²/V s. For vertical fields of order 10 V/m, this mobility fails to produce kilometer-scale electron drifts on time scales of order 1 ms.

Two alternative mechanisms actually do contribute significantly to the depletion. One is the local rarefaction of the electron population due to the upward flow of electrons over short distances in the presence of a strong vertical mobility gradient. As discussed above, the electron mobility grows exponentially with altitude due to the declining neutral density. The decade scale height of this variation is 16 km, from Figure 2.2. The effect of this increasing mobility on the electron flux is magnified by the fact that the total electron density is itself increasing with altitude. By contrast, the decline of the vertical electric field between 40 and 60 km altitude is much more gradual, with a decade scale height of 27 km, from Figure 4.1. The net effect is a strong divergence in the electron flux over these altitudes, and a corresponding decrease in density. The
Figure 4.5: Change in electron density from quiescent values for conditions of Figure 4.1. (a) Change in density throughout simulation volume. The large depletion is due primarily to dissociative attachment. The enhancements are due to transport of electrons. (b) Change in electron density along simulation axis. Negative values are indicated by broken curve, positive values by solid curve.
fractional density decrease that results from this mechanism is 0.2 % per ms. While significant, this mechanism is by itself unable to account for the magnitude of the density depletion.

It turns out that the bulk of the electron density depletion observed in Figure 4.5 is the result of local dissociative attachment of electrons to molecular oxygen, due to the elevated electron temperature. This process has a sharp threshold near 6 eV [Schulz, 1962] and comes into play at much lower electron temperatures than impact ionization. Rates of dissociative attachment are compared to those of impact ionization in Table 4.1.

<table>
<thead>
<tr>
<th>Electron Temperature (eV)</th>
<th>Dissociative Attachment Rate (s⁻¹)</th>
<th>Impact Ionization Rate (s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.216E+04</td>
<td>0.99</td>
</tr>
<tr>
<td>0.7</td>
<td>0.255E+03</td>
<td>0.162E-02</td>
</tr>
<tr>
<td>0.4</td>
<td>0.115E+01</td>
<td>0.107E-08</td>
</tr>
</tbody>
</table>

Table 4.1 Rates of dissociative attachment and impact ionization for several temperatures at an altitude of 60 km. Rates are per electron, per second.

It is clear from the table not only that dissociative attachment is vastly more important than impact ionization at the temperatures and altitudes under consideration, but also that dissociative attachment can easily account for teas of percent electron density decrease per ms at the 0.7 eV electron temperature. Moreover, the rate of dissociative attachment increases so rapidly with temperature that even a minor increase in temperature can produce a ten-fold increase in the rate.

The actual rate at 0.7 eV implies a density depletion notably more intense than the one observed, but examining this one rate in isolation is meant only to indicate that dissociative attachment is capable of producing such a depletion. In practice, the effects of dissociative attachment are moderated by an increase in electron detachment and the net result can be obtained accurately only from a solution of the chemical rate equations (2.110). Figure 4.5 incorporates that solution.

In addition to the one major density depletion, Figure 4.5 also reveals a density enhancement at large radii near 60 km altitude and another close to the axis at even higher altitudes. Electron transport is responsible for both of these features. First, the decrease with radius of the radial electric field magnitude in the presence of relatively constant mobilities in the radial dimension produces a convergence of electron flux at high altitudes and large radii. The result is enhanced electron density in this region. The slight electron density enhancement at the highest altitudes, near the simulation axis is the result of bulk electron transport in the small upward fields at those high altitudes.
4.1.2 Temporal Evolution of Simulation System

We have so far focused on the conditions that are obtained at a particular instant in time, part way through the lightning discharge. We now turn to the evolution of the system before, during, and after the peak lightning current.

Figure 4.6 shows the electron temperature and a related quantity, the total thermal energy content of a high altitude control volume, as a function of time. The results are taken for a point on the simulation axis, at an altitude of 56.4 km. The lightning stroke responsible for the temperature variation is the same as the one used in the preceding section, namely, a positive CG discharge that lowers 200 C to ground with an e-folding time constant of 1.25 ms.

Figure 4.6a divides the time span of the simulation into the electrification, discharge, and relaxation periods, and details the development in each period. The temperature during the electrification period increases modestly and gradually from a quiescent value of 0.025 eV to a steady value of approximately 0.1 eV after several seconds. The modest nature of this increase results from the fact that on the slow time scale of the electrification, the medium at the intervening altitudes is able to largely screen out the electric field originating at cloud altitudes.

By contrast, the rapid cloud height charge removal during the discharge outpaces the low altitude plasma response, and the electron temperature at 56.4 km grows dramatically. The explosive nature of the temperature increase is emphasized by Figure 4.6b, which shows temperature as a function of simulation time in seconds. The temperature increase saturates at .72 eV after 3.4 ms of discharge and then begins to relax at a rate that is roughly consistent with the local dielectric relaxation rate of the medium.

The total thermal energy content of the control volume displays an evolution quite similar to that of the temperature. The control volume extends from 40 to 70 km in altitude, and out to a radius of 20 km. The maximum thermal energy is 934 J, at 0.9 ms after discharge onset. One might anticipate the evolution of the total energy to be smoother by virtue of the averaging effect of the integration. In fact, because the integrand is weighted by the electron density, which increases rapidly with altitude, the total energy is most strongly influenced by the highest altitude slice of the control volume, and is dominated by variations in the electron temperature there.

The temporal behavior of the electric field as a function of altitude is examined in Figure 4.7. The basic picture at altitudes above 40 km is that the electric field remains essentially zero.
Figure 4.6: Time history of thermal quantities associated with positive CG discharge that lowers 200 C to ground from 15 km altitude, with peak current of 160 kA. (a) Electron temperature at 56.4 km altitude on the simulation axis, displayed in terms of simulation steps. (b) Electron temperature displayed versus simulation time. (c) Total thermal energy of electrons in control volume between 40 and 70 km altitude and 0 to 20 km radius.
Figure 4.7: Altitude dependence of lightning-induced electric field pulses. Shapes of electric field pulses at altitudes of (a) 72.5 km, (b) 59.1 km, (c) 45.7 km, and (d) 32.2 km. Peak field values and pulse widths are given in Table 4.2. (e) Electric field pulse width $\tau_{Ez}$, as a function of altitude, compared to local dielectric relaxation time, $\tau_D$, and chemical relaxation rates. Chemical relaxation rates are the decay times of a positive increment in the density of a single charged species away from equilibrium. Species labels $n_c$, $n_p$, $n_m$, and $n_e$ refer to electrons, small positive ions, negative ions, and cluster ions, respectively.
through the electrification period and then exhibits a negative going pulse the commences with the onset of the lightning discharge. The absence of an electric field during cloud electrification is once again the result of the dielectric response of the medium at lower altitudes keeping pace with the relatively slow growth in cloud charge. The magnitude of the field pulse decreases rapidly with altitude, and appears at slightly earlier times at higher altitudes, as shown in Table 4.2.

<table>
<thead>
<tr>
<th>Altitude (km)</th>
<th>Pulse Magnitude (V/m)</th>
<th>Time of Peak (ms)</th>
<th>Decay Time, Half Max (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>72.5</td>
<td>-13.0</td>
<td>0.57</td>
<td>1.9</td>
</tr>
<tr>
<td>59.1</td>
<td>-278.0</td>
<td>2.6</td>
<td>4.0</td>
</tr>
<tr>
<td>45.7</td>
<td>-1160.0</td>
<td>7.8</td>
<td>67.2</td>
</tr>
<tr>
<td>32.2</td>
<td>-5419.0</td>
<td>162.0</td>
<td>134.0</td>
</tr>
</tbody>
</table>

Table 4.2 Characteristics of high altitude electric field pulses at several altitudes.

The declining magnitude of the pulses with altitude results directly from the spatial decay of the instantaneous electric field produced at low altitudes by the sudden removal of charge at cloud height. This spatial decay is due to a combination of $1/r^2$ free space decline and dielectric shielding by the intervening medium. The decrease in pulse width and the earlier appearance of the peak are both linked to the increasing speed of the dielectric response with altitude. Pulse width decreases as higher total conductivity permits more rapid charge redistribution in response to the applied fields. The peak electric field appears earlier because at the highest altitudes the shielding response of the medium is well underway even as the pulse is developing.

The concept of a local dielectric response by the medium to fields originating at low altitudes has wide application in explaining the temporal behavior of the electric field observed at high altitudes. The basis for this response is found in current continuity (2.125), Ohm’s Law (2.140), and Poisson’s Equation (2.20). We reproduce these equations here under the simplifying assumption that there is no charge source.

$$\nabla \cdot J + \frac{\partial \rho}{\partial t} = 0 \quad (4.2)$$

$$\bar{J} = \sigma \bar{E} \quad (4.3)$$

$$\nabla \cdot (\varepsilon \bar{E}) = \rho \quad (4.4)$$

The physical scenario embodied by these equations is that the existing charge distribution establishes an electric field according to Poisson’s Equation. This electric field causes currents according to Ohm’s Law. These currents redistribute charge according to current continuity, with
the overall result that the new charge distribution is itself reduced in magnitude, and results in reduced electric fields. If the total conductivity and the dielectric constant are uniform in space, then we may combine equations (4.2), (4.3), and (4.4) to write:

\[
\frac{1}{\sigma} \frac{\partial \rho}{\partial t} + \frac{\rho}{\varepsilon} = 0 \Rightarrow \rho(\vec{r}, t) = \rho_0(\vec{r}) e^{-t/\tau_D}
\]  

(4.5)

where \(\rho_0(\vec{r})\) is the initial charge distribution, and \(\tau_D = \varepsilon_0/\sigma\) is known as the dielectric relaxation time. We use the permittivity of free space because we are dealing with free charges in a vacuum and not bound charge. Thus, a homogeneous medium will respond so as to smooth out the charge distribution, and extinguish electric fields on a time scale of the order of the dielectric relaxation time. \(\tau_D\) is the characteristic time of the medium's local response to an applied electric field.

We compare the local dielectric relaxation time to the lightning-induced electric field pulse width in Figure 4.7e. The pulse width follows the dielectric relaxation time remarkably well over an altitude range extending from 40 km to 70 km. This indicates that the pulse width is dominated by the local plasma dielectric response at these altitudes. Regardless of the evolution of fields and charges at lower altitudes, the local medium redistributes charge so as to screen out the electric field on a time scale characteristic of its own response. At still higher altitudes, the pulse width departs from \(\tau_D\) because the high electron mobility makes the response of the medium significantly nonlocal.

At altitudes below 40 km, the decay of the electric field pulse is dominated by processes other than the local dielectric response. This is demonstrated quite dramatically by Figure 4.7e, which shows that the pulse width departs from the dielectric response time at low altitudes, and instead follows the chemical relaxation time of the negative ion density. The four solid curves in the figure, corresponding to each of the charged species, show the duration required for an initial increment in any one species to decay away due to the chemical effects modeled in Section 2.6. In most instances, an increment in only one species is not actually observed, the chemical evolution instead being a complicated interplay among several species. However, the situation that exists at the start of the lightning discharge is dominated by an excess in small negative ions over the equilibrium population. This excess develops as electrons are drawn into the region during the electrification period and then rapidly converted to small negative ions. It is seen from Figure 4.7e that three-body attachment results in extremely rapid decay of the electron population at altitudes
below 40 km. The outcome of the loss of electrons is a corresponding growth in the negative ion population.

This situation persists for as long as the positive electric field is growing during the electrification period. With the onset of the lightning discharge, the downward flow of electrons ceases. In fact, an electron depletion quickly develops, as was shown in Figure 4.5. The negative ions, on the other hand, are immobile on millisecond time scales and remain in place until they are removed by the chemical processes of recombination and electron detachment. In regions where the local dielectric response takes longer than the chemical absorption of the negative ions, the electric field disturbance due to the discharge persists for as long as the ions present net space charge.

This is precisely the scenario that unfolds at low altitudes. As seen in Figure 4.7c, the local dielectric response takes longer and longer with decreasing altitude due to the growing neutral density. This trend reverses below 22 km altitude because of the great density of small ions and electrons that accumulate there during the electrification process. In any case, between 20 and 40 km altitude, the dielectric response time exceeds the chemical relaxation time of the negative ions. As a result the electric field pulse width at these altitudes is governed not by the dielectric response of the medium, but by the chemical response to the excess negative ion population.

4.1.3 Applicability of Simulation Approach

The large peak values of electric field encountered in the simulation results of Sections 4.1.1 and 4.1.2 cause us to re-examine several of the assumptions that entered into the development of our fluid model. In particular, we ask whether the fundamental assumption of an isotropic electron velocity distribution holds in the presence of the fast electron drifts caused by such large peak fields. Moreover, our simplification of the fluid equations was based upon parameter values that depended in part upon assumed electric fields. We should therefore revisit our simplifications.

In developing the fluid equations, we examined conditions at two benchmark altitudes, 30 km and 70 km, in Table 2.7. The anticipated magnitudes of the electric field at these altitudes were 100 V/m and 1 V/m, respectively. In practice, we encountered peak magnitudes of order 1000 V/m and 10 V/m. Table 4.3 compares the resulting electron drift speeds to the mean thermal speed at the conservative electron temperature of 0.5 eV.
<table>
<thead>
<tr>
<th>Altitude (km)</th>
<th>E-field (V/m)</th>
<th>T_e (eV)</th>
<th>Mobility (m²/Vs)</th>
<th>Drift Speed (m/s)</th>
<th>Thermal Speed (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>10</td>
<td>0.5</td>
<td>8.9E+03</td>
<td>8.9E+04</td>
<td>4.2E+05</td>
</tr>
<tr>
<td>30</td>
<td>1000</td>
<td>0.5</td>
<td>2.9E+01</td>
<td>2.9E+05</td>
<td>4.2E+05</td>
</tr>
</tbody>
</table>

Table 4.3 Electron drift speed versus thermal speed for conditions obtained during the simulation.

It is clear that while the drift speed does not exceed the mean thermal speed, it comes within a factor of two, particularly at low altitudes, where the fields are most intense. Under such circumstances, the approximation of the electron distribution function as isotropic is at best adequate, and we must recognize that we are about at the limit of the applicability of that approximation. The near-breakdown of the isotropy assumption does not affect the fluid formulation in general, for it admits arbitrary particle distributions, and accounts for the anisotropy as a drift. However, our calculation of collisional quantities, based both on the Maxwellian and the Druyvestein velocity distributions, did assume isotropy.

By computing the collisional quantities using an isotropic distribution, we in effect ignore the contribution of the electron drift kinetic energy to the mean electron speed. We can estimate the magnitude of this effect by adding the electron drift to the mean thermal speed in root-mean-squared fashion. The resultant mean speed is 4.29E+05 m/s at 70 km altitude, and 5.02E+05 m/s at 30 km. These values represent rather modest increases in the mean electron speed from Table 4.3. As a result, we conclude that the neglect of electron drifts has produced a correspondingly modest underestimate of collision rates. The ultimate effect is a minor overestimate of the electron temperature on the order of several percent.

We are, in addition, concerned about the effect of large representative electric field magnitudes on the parameters in our nondimensionalized fluid equations, as listed in Table 2.6. Referring to that table, we see that larger than expected values of $E_0$ lead to increased values of $\alpha$, and decreased values of both $\beta$ and $\eta$. We make no simplification based upon the value of $\alpha$, so changes in it do not worry us. The same is true of $\beta$, although we can point out that reducing the magnitude of $\beta$ from its already small values enhances the validity of Ohm's Law. Finally, the reduction in $\eta$ from values that are much greater than 1 to values of order 1 implies that we cannot ignore the electron drift kinetic energy in comparison to the electron thermal energy, as has been done in deriving the simplified energy transport equation (2.161). This is precisely the same issue.
that was addressed in the preceding paragraph, and the consequences of the action are equally minor.

4.2 Comparison to Physical Observations

We have accorded special attention to our simulation results regarding elevated electron temperatures and high altitude electric field transients. Our interest stems from the fact that these are two macroscopic variables for which observations are available for comparison. As noted in Section 1.1.1, *in-situ* measurements at high altitudes above thunderstorms coincident with lightning observations are difficult to come by, as they generally require balloons or sounding rockets as measurement platforms. We are thus limited to a rather brief comparison of our simulation results with physical observations.

The data that is perhaps best suited for comparison to our simulation findings consists of electron temperatures inferred from optical measurements of sprites [Green et al., 1996]. The unique suitability of the Green et al. [1996] measurements arises from the fact that they are guaranteed to be associated with strong positive cloud-to-ground lightning discharges of the type we have been modeling. The authors analyze spectral data from a sprite column near 70 km altitude to obtain the relative population of the vibrational states of the ambient nitrogen molecules. From these results they infer the temperature of the Maxwellian velocity distribution of the electrons responsible for exciting the observed emissions. They conclude that the required electron temperature is in the range of 0.4 to 2.0 eV, which is in good agreement with our predictions.

The simulation predicts elevated electron temperatures over an extended altitude range from 40 to 70 km, with a peak at the highest altitudes. Sprite observations report the most intense optical emissions from the much narrower range of 66-74 km altitude [Sentman et al., 1995]. The dominance of a narrow altitude range arises because the amount of luminosity depends not only on the electron temperature, but also upon the electron density. The key determinant of the rate of collisional excitation of the ambient neutrals is the density of electrons in regions where the applicable threshold in electron temperature is met. In other words, regardless of the degree of heating of the electrons at lower altitudes, there are insufficient numbers of electrons present there to perform significant excitation. It may thus be concluded that not only the magnitude, but also the gross morphology of the electron heating predicted by the simulation, is in agreement with sprite observations.
The comparison of electric field transients predicted by the simulation to those that are observed is hampered by a lack of knowledge regarding the causative lightning strike. For example, Blakeslee, et al. [1989] reports balloon-borne measurements of DC electric fields above an active thunderstorm. Electric field magnitudes exceeding 5 kV/m are observed regularly at an altitude of 20 km above a thunderstorm. Field discontinuities of between 1 and 4 kV/m are observed to coincide with lightning discharges. Unfortunately, the amount of charge lowered to ground is not known for the associated lightning discharges, and the polarity of the lightning is generally negative. Nevertheless, we may run the simulation for an average negative cloud-to-ground lightning discharge, as defined by Table 1.2. The result is a 1.3 kV/m electric field transient at 20 km altitude. This is in rough agreement with the electric field discontinuities observed by Blakeslee, et al. [1989].

Electric field transients have been observed at ionospheric altitudes over active thunderstorms, as discussed in Section 1.1.1. Since only the most intense examples of lightning produce measurable transients at such high altitudes [Kelley et al., 1985], our simulation results for a relatively intense, positive cloud-to-ground discharge should be quite applicable to these observations. However, most of these rocket- and satellite-borne measurements have been taken at altitudes above the domain of the simulation. One measurement that is directly comparable to the simulation results was taken by the Thunder Lo sounding rocket at an altitude of 88 km over the Atlantic off the Maryland coast. The measurements revealed transients of 30 mV/m amplitude, lasting for 10 ms or more. This is both larger and longer than the transients predicted by the present simulation. According to the simulation, we should expect transients of 1.6 mV/m amplitude and 1 ms duration at an altitude of 85 km for the lightning stroke of Figure 4.1. At these great altitudes, Kelley et al. [1985] measured a significant electromagnetic component of the lightning-induced disturbance. This would account for the much higher amplitude of the observation than the simulation result, which treats only the electrostatic component. Moreover, Kelley et al. [1990] showed that significant dispersion of the lightning-induced electromagnetic pulse takes place as the pulse propagates. This would account for the longer pulse duration.

4.3 Influence of Parameters

The simulation output depends on a great many parameters associated with the ambient environment and with the lightning discharge itself. Many of these, such as the reaction rates and the neutral density, have fixed, known profiles versus altitude. We have modeled these parameters
carefully, as described in Chapter 2. Moreover, we have examined the sensitivity of the individual models and the overall simulation to such parameters. We have found that varying these parameters within the tolerance to which they are known causes no disproportionate effect on the simulation output. Thus, we treat this class of parameters as fixed and known.

Other parameters are, in fact, environmental characteristics that may vary greatly with the passage of time from one thunderstorm to the next. Still others are characteristics of the lightning that may vary from one stroke to the next. It is the physical impact of this class of parameters that concerns us in this section. We consider the effects on the simulation of one environmental characteristic, the high altitude quiescent electron density profile; and two lightning characteristics, total charge lowered to ground and stroke duration.

### 4.3.1 High Altitude Electron Density

As discussed in Section 2.2.2, the high altitude electron density is extremely variable below the ledge at 80 km altitude. Among the sources of this variability are the diurnal variation of the solar zenith angle, variations in geomagnetic location, and short-term disturbances in solar activity [McNamara, 1985]. These influences are beyond the scope of our model. According to the sources cited in Section 2.2.2, the possible electron density profiles can be characterized by a range of scale heights. We examine four possible cases within the range: decade scale heights of 2.5 km, 5.0 km, 7.5 km, and 10.0 km.

The shape of the altitude profile of electron density in each of these cases is displayed in Figure 4.8a. The case of the 2.5 km scale height corresponds to the most tenuous profile, while the 10.0 km scale height corresponds to the most dense profile. The other two cases appear monotonically in between. The difference among the density profiles is confined to the altitude range of 40 to 80 km. Our test lightning discharge for the purposes of this study is a positive CG discharge that lowers 200 C to ground from a monopolar cloud at 15 km altitude. The peak current is 160 kA.

The principal effect of varying the ambient electron density at high altitudes is to change the conductivity of the medium. The larger values of conductivity associated with the denser electron profiles extinguishes the quasi-electrostatic field at lower altitudes. As a result, electron heating is confined to somewhat lower altitudes, as demonstrated in Figure 4.8b, but the altitude of the maximum electron temperature is not affected.
Figure 4.8: Influence of electron density profile on simulation results. In all panels, the broken curve corresponds to the results for the most tenuous density profile. The other curves, in order of increasing separation from the broken curve, correspond to more and more dense profiles. (a) Test profiles of electron density, parameterized by the decade scale height of density decrease with altitude below 80 km. Scale heights used are 2.5 km, 5.0 km, 7.5 km, and 10.0 km. 2.5 km represents the most tenuous profile, and 10.0 km, the most dense. (b) Electron temperature versus altitude along simulation axis at 3.0 ms after onset of positive cloud-to-ground discharge that lowers 200 C of charge to ground, with peak current of 160 kA. (c) Local dielectric relaxation time versus altitude along simulation axis, during electrification period. (d) Time series of electron temperature at 56.4 km altitude on simulation axis. Shows complete lack of heating during electrification period for cases of dense ionosphere.
A direct consequence of enhanced conductivity is the reduction of the dielectric relaxation
time, shown in Figure 4.8c. We saw in Section 4.1.2 that the local dielectric relaxation time is a
relevant concept at altitudes above 40 km. In the cases shown, the denser profiles yield shorter
relaxation times above 40 km. Two effects arise from reduced values of $\tau_D$, shorter electric field
transient pulse widths and diminished heating during electrification. According to Figure 4.8d,
only in the cases of the two most tenuous electron density profiles is any heating above neutral
temperature observed. This indicates that under dense plasma conditions, the local medium at 56.4
km altitude is effectively shielded from the electric fields originating from the slowly growing cloud
charge. The electron temperatures during the discharge period are not affected.

Table 4.4 calls out the variation of three specific quantities with the high altitude electron
density scale height.

<table>
<thead>
<tr>
<th>$n_e$ Scale Height (km)</th>
<th>$E_z$ pulse amplitude (V/m)</th>
<th>Maximum $T_e$ (eV)</th>
<th>Maximum $\Delta n_e$ (m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>most tenuous</td>
<td>-20.2</td>
<td>0.71</td>
</tr>
<tr>
<td>5.0</td>
<td></td>
<td>-18.7</td>
<td>0.71</td>
</tr>
<tr>
<td>7.5</td>
<td></td>
<td>-19.8</td>
<td>0.73</td>
</tr>
<tr>
<td>10.0</td>
<td>most dense</td>
<td>-20.6</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Table 4.4: Variation of vertical electric field pulse amplitude, maximum electron
temperature $T_e$, and maximum electron density depletion $\Delta n_e$ with electron density profile. $E_z$ is
measured on the simulation axis at 70 km altitude. Maximum electron temperature occurs at 56.4
km in all cases.

We note that while maximum temperature and field pulse amplitude are not influenced by the
electron profile, the peak absolute electron density depletion shows an unmistakable trend of
growth with increasing plasma density. This results from the fact, pointed out in Section 4.1.1,
that the density depletion results from a temperature related enhancement in the dissociative
attachment of electrons to oxygen. Thus, with more electrons present at high altitudes in the case
of the denser profiles, the rate of dissociative attachment is increased, explaining the greater density
depletion.

Finally, we should point out that all of the simulation results presented in this work, with
the exception of those in this section, assume a high altitude electron density profile consistent with
the 2.5 km scale height.
4.3.2 Magnitude of Cloud Charge Lowered to Ground

In our discussion of lightning phenomenology in Section 1.1.2, we found that the total charge lowered to ground during a positive cloud-to-ground discharge can vary widely. Five percent of positive strokes lower less than 20 C and five percent lower more than 350 C, and the median value is 80 C [Uman, 1987, p. 124]. The actual value for any given stroke is unpredictable and liable to vary from stroke to stroke within the same storm, depending upon the intimate details of the cloud electrification process. In principle, sprite observations could be correlated with lightning charge using data from the National Lightning Detection Network (NLDN) [Orville et al., 1983]. In anticipation of such a study, we examine the predicted relationship between the magnitude of cloud charge lowered to ground, and electron heating in the overlying medium.

Figure 4.9a shows the electron temperature versus altitude profiles that result 3 ms after the onset of four different lightning discharges. The discharges differ only in the total charge lowered to ground, with values of 20 C, 50 C, 100 C, and 200 C. In each case the cloud charge is centered at an altitude of 15 km, and the discharge has an e-folding time of 1.25 ms. As a result of fixing the discharge duration, the lightning current varies among the cases considered, in proportion to the total charge, with values of 16 kA, 40 kA, 80 kA, and 160 kA, respectively. We may thereby be combining the effects varying of total charge with those of varying peak lightning current. It is nevertheless appropriate to proceed in this manner because lightning stroke duration is an independent parameter [Uman, 1987, p. 188].

We see from the figure that as lightning charge increases, the temperature rises at all altitudes in such a way that the shape of the temperature profile does not change significantly. The corresponding electron density depletions remain confined to the altitude range between 40 and 70 km in all cases, as shown in Figure 4.9b. The peak of the density depletion simply becomes sharper and more intense with increasing lightning charge. In a similar vein, Figure 4.9c shows that the only change in the time variation of the electron temperature is due to the increase in peak temperature with lightning charge. We note that heating during the electrification period also grows with lightning charge.

The juxtaposition of Figures 4.9a and 4.9b illustrates that the peak of the electron density depletion always occurs at slightly higher altitude than the peak temperature (60.4 km versus 56.4 km, for the example of 200 C lightning charge). This is because the background electron density is
Figure 4.9: Influence of cloud charge lowered to ground by lightning stroke. Broken curve shows results for 200 C in all panels. Results for 100 C, 50 C, and 25 C appear with increasing separation from the broken curve. In all cases, discharge originates at 15 km altitude, with an e-folding decay time of 1.25 ms. (a) Electron temperature versus altitude along simulation axis at 3 ms after onset of discharge. (b) Change in electron density relative to initial conditions at locations along simulation axis at 3 ms after onset of discharge. (c) Time series of electron temperature at 56.4 km altitude on simulation axis.
rising rapidly with altitude in the region, so that the maximum rate of dissociative attachment, which occurs in the vicinity of the peak temperatures, is skewed to slightly higher altitudes.

As a whole, these results reflect a general, monotonic intensification of the system response due to a pure increase in the magnitude of the lightning stimulus. Table 4.5 puts specific numbers to this trend.

<table>
<thead>
<tr>
<th>Lightning Charge (C)</th>
<th>$E_z$ pulse amplitude (V/m)</th>
<th>Maximum $T_e$ (eV)</th>
<th>Maximum $\Delta n_e$ (m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>-20.2</td>
<td>0.71</td>
<td>-1.82E+06</td>
</tr>
<tr>
<td>100</td>
<td>-8.8</td>
<td>0.51</td>
<td>-1.03E+06</td>
</tr>
<tr>
<td>50</td>
<td>-4.7</td>
<td>0.40</td>
<td>-5.79E+05</td>
</tr>
<tr>
<td>20</td>
<td>-1.4</td>
<td>0.32</td>
<td>-2.48E+05</td>
</tr>
</tbody>
</table>

Table 4.5: Variation of simulation results with total cloud charge lowered to ground. Vertical electric field pulse amplitude, $E_z$, is measured at an altitude of 70 km, on simulation axis.

According to the table, the amplitude of the high altitude electric field pulse grows faster than linearly with the lightning charge, while the maximum electron temperature grows slower than linearly. The relatively greater effect on the electric field stems from the reduced high altitude conductivity that accompanies the higher electron temperatures. As pointed out in Section 4.1, lower conductivity permits the instantaneous electric field to penetrate to higher altitudes than would otherwise be possible. On the other hand, the effects of increasing lightning stimulus on peak electron temperature are mitigated by the rapidly increasing rate of inelastic collisional energy loss with increasing temperature. Thus, a given increment in lightning charge results in a less than proportional increment in peak electron temperature.

### 4.3.3 Lightning Stroke Duration

Positive cloud-to-ground lightning is typically characterized by a period of continuing current that can extend the duration of the return stroke by any amount from several milliseconds to over a hundred milliseconds [Fuquay, 1982]. We compare the results of four lightning discharges, each with a different $e$-folding decay time, hence a different duration. We define the duration to be four times the decay time. The examples considered have durations of 20 ms, 10 ms, 5 ms and 2 ms. In each case, the same total charge is lowered to ground, 200 C, from a charge distribution centered at 15 km altitude. Holding the cloud charge fixed in the face of varying stroke duration results in differing lightning current for each case. In fact, it is the magnitude of the
lightning current that appears to have the dominant effect on our results. The peak values of lightning current associated with each of our examples are 40 kA, 80 kA, 160 kA, and 400 kA, respectively.

In Figure 4.10a, we see that changes in the stroke duration produce significant effects on both the maximum electron temperature and the shape of the high altitude temperature profile. Longer strokes lead to lower maximum temperatures and a quenching of electron heating in an altitude range that grows downward with increasing stroke duration. Because these differences are evident at a fixed time after the onset of discharge, we know that they are not the result of increased stroke duration per se, but instead stem from the slower rate of charge removal associated with the lower initial lightning currents of the longer strokes.

The pronounced quenching of electron heating in a region of varying altitude extent is a consequence of the relative time scales of the discharge and the dielectric response of the medium. In all cases, at a sufficiently high altitude the dielectric response of the medium defeats the electric field growth associated with cloud height charge removal. This is true down to lower and lower altitudes as stroke duration increases—in accordance with the lengthening of the dielectric response time at lower altitudes, as shown in Figure 4.7e. As a result, electron heating is quenched in an altitude band that extends to lower and lower altitudes with increasing stroke duration.

The decrease in maximum electron temperature with increasing stroke duration is a direct effect of the varying altitude extent of the quenched region. The value of the temperature at a given location is not significantly different when the maximum is attained there from when the maximum is at a higher altitude. In these examples, the variation of the maximum temperature with stroke duration is mainly a characteristic of the instantaneous variation in electron temperature with altitude that stems from the altitude profile of total conductivity, as discussed in Section 4.1.1.

Figure 4.10b shows that a lowering of the altitude of peak electron depletion accompanies the longer duration lightning strokes. The location of the peak depletion follows the location of the temperature maximum, for that is where the highest rates of dissociative attachment are obtained. The changes in altitude of the maxima in both quantities are highlighted in Table 4.6.
Figure 4.10: Influence of lightning stroke duration on simulation results. Broken curve shows results for 20 ms discharge in all panels. Results for 10 ms, 5 ms, and 2 ms appear with increasing separation from the broken curve. All four discharges originate at 15 km altitude and lower 200 C of cloud charge to ground. (a) Electron temperature versus altitude along simulation axis at 3 ms after onset of discharge. (b) Change in electron density relative to initial conditions at locations along simulation axis at 3 ms after onset of discharge. (c) Time series of electron temperature at 56.4 km altitude on simulation axis.
<table>
<thead>
<tr>
<th>Stroke Duration (ms)</th>
<th>$E_z$ pulse amplitude (V/m)</th>
<th>Maximum $T_e$ (eV)</th>
<th>Altitude of Maximum $T_e$ (km)</th>
<th>Maximum $\Delta n_e$ (m$^3$)</th>
<th>Altitude of Maximum $\Delta n_e$ (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-34.6</td>
<td>0.77</td>
<td>60.4</td>
<td>-1.21E+06</td>
<td>65.8</td>
</tr>
<tr>
<td>5</td>
<td>-20.2</td>
<td>0.71</td>
<td>56.4</td>
<td>-1.82E+06</td>
<td>60.4</td>
</tr>
<tr>
<td>10</td>
<td>-12.0</td>
<td>0.66</td>
<td>53.7</td>
<td>-2.50E+06</td>
<td>57.8</td>
</tr>
<tr>
<td>20</td>
<td>-8.4</td>
<td>0.62</td>
<td>49.7</td>
<td>-3.07E+06</td>
<td>55.1</td>
</tr>
</tbody>
</table>

Table 4.6: Variation of simulation results with lightning stroke duration. Vertical electric field pulse amplitude, $E_z$, is measured at an altitude of 70 km, on the simulation axis. Maximum electron temperature, $T_e$, and maximum electron density depletion, $\Delta n_e$, are taken at 3 ms after onset of discharge.

We note that the intensity of the depletion increases with longer duration strokes, in spite of the fact that the associated electron temperature is smaller in those cases. This is explained by the rapidly growing neutral densities that are encountered as maximum electron temperature occurs at lower altitudes.

Figure 4.10c shows that the different cases exhibit a significant variation in electron temperature during the early discharge period, with the shorter discharges producing higher temperatures. In spite of this variation, the electron temperature relaxes to a common value at 10 ms after onset of discharge, regardless of stroke duration. This confirms that the effects attributed to stroke duration have at their root the associated differences in lightning current.

4.4 Special Phenomena

We now present several phenomena predicted by the simulation. These phenomena will be important for one of two reasons. On the one hand, they may contribute to explaining observed sprite features and correlations. On the other, they may point the way to as yet unobserved effects. The phenomenon of high altitude charge separation, for example, contributes to an explanation of the strong correlation observed between sprites and positive, not negative, cloud-to-ground discharges. Equally important are the persistent effects of ion relocation due to a lightning discharge, which have bearing on the longer-lived, sprite-related features known as blue jets, as well as on subsequent lightning discharges. Finally, a highly localized temperature depression is found to move rapidly downward during the discharge period, at a speed consistent with the lower boundary of elves observed by Fukunishi [1996].
One additional phenomenon predicted by the simulation has not yet been observed. It is a shock-like structure in the electron density that propagates slowly downward during the relaxation period. The reason that this phenomenon has not been observed is that it could only be detected by time correlated, in-situ observations or active remote sensing techniques that have not yet been employed.

4.4.1 High Altitude Charge Separation

The presence of space charge in the 20 km to 40 km altitude region, and the effects that result from it, become apparent when we compare simulations of positive versus negative cloud-to-ground discharges. We make a comparison between two lightning strokes of equal 200 C magnitude, but opposite sign—each depleting a charge center at 15 km altitude. In both cases, the electrification is exponential, lasting 5 s (1.25 s time constant), followed by a 5 ms exponential discharge (1.25 ms time constant) and a 1 s relaxation period.

The simulation shows a marked difference in peak electron temperature and total energy imparted to the high-altitude volume when comparing the results of positive and negative CG discharges. Figure 4.11 compares the electron temperature obtained for the two cases, everywhere in the simulation volume, 3 ms after onset of the discharge. The maximum temperature observed for the positive discharge is roughly twice that for the negative discharge. Temperature versus altitude profiles along the simulation axis are shown in Figure 4.12a. The positive CG discharge exhibits a peak temperature of 0.73 eV at an altitude of 72.5 km, while the negative CG discharge shows 0.4 eV at 71.2 km.

By changing nothing but the polarity of the stroke between the two cases, we control for other parameters that typically distinguish positive from negative CG discharges, and thereby isolate the effect of polarity. This is significant because we saw in Section 1.1.2 that positive and negative discharges differ greatly in attributes other than polarity. One of the chief differences between the two types of discharge is the typical altitude of the associated charge centers, namely 7 km for negative CG and 9 - 13 km for positive CG in a summer thunderstorm [Krehbiel, 1986]. In addition, the magnitude of the charge lowered to ground also differs greatly between the two types, with a median value of 80 C for positive flashes and only 7.5 C for negative flashes [Uman, 1987, p. 124].
Figure 4.11: Electron temperatures throughout the simulation volume at 3 ms after onset of lightning discharge. (a) +200 C lowered to ground from charge center at 15 km altitude on axis. Greater off-axis heating and faster radial falloff at high altitude—relative to part b—is the result of high altitude charge separation not produced by negative cloud-to-ground discharge. (b) -200 C lowered to ground from charge center at 15 km altitude on axis.
Figure 4.12. Effects of positive versus negative cloud-to-ground (CG) lightning. ± 200 C lowered to ground from 15 km altitude. Altitude profiles of (a) electron temperature, (b) absolute charge density, and (c) total conductivity on simulation axis 1.5 ms after onset of discharge. Broken curves show results of +200 C lowered to ground from 10 km altitude. Charge density for +CG discharges is negative. Temporal evolution of (d) vertical electric field and (e) charge density at 34.9 km altitude on simulation axis.
Due to the magnitude of the disparity in these characteristics, we are led to assess
their relative importance. Figure 4.12a addresses the effect of cloud charge polarity by showing
the altitude profile of temperature for a positive CG discharge from 10 km altitude. The effect of
changing polarity is greater than that of reducing the altitude of the positive charge center from 15
km to 10 km. Peak temperature for the positive CG discharge from 10 km is 0.56 eV—midway
between the results for the positive and negative discharges from 15 km. Although not shown in
the figure, the combined effect of reducing the lightning charge and lowering the altitude of the
charge center is very strong; a positive CG discharge that lowers 20°C from an altitude of 10 km
produces peak electron temperature of only 0.28 eV, less than even the negative CG discharge
shown in Figure 4.12a.

Of the three differences between positive and negative CG lightning discharges—polarity,
altitude, and magnitude—the magnitude of the charge lowered to ground has the strongest effect on
the intensity of mesospheric electron heating, followed by polarity and then altitude. However,
positive CG discharges involve a distinct mechanism for enhancing electron heating, namely charge
separation, that does not come into play for negative discharges. It is the phenomenon of charge
separation that accounts for the effect of polarity.

Figure 4.12b displays evidence of large negative charge densities between 20 and 40 km
altitudes, for both cases of positive CG discharge shown. Space charge is virtually absent from
this region in the case of the negative CG discharge. We see from Figure 4.12e that the charge
density produced by the positive lightning, at the representative altitude of 34.9 km, evolves over a
period ~ 1 s during the electrification of the thundercloud. It dissipates over a similar interval in
the relaxation period. The fact that the space charge shown develops during the electrification
period indicates that it is not simply residual neutralizing charge that builds up during the
electrification and is uncovered only during the rapid charge removal of the discharge period. The
mechanism responsible for the space charge is the macroscopic flow of ions and electrons in the
presence of strong gradients in both neutral and electron densities.

Charge introduced at cloud height during the electrification is neither fully screened nor
fully neutralized by the low altitude medium—even on the long time scale of electrification period.
As a result, finite, positive (upward) electric fields develop above the cloud at all altitudes up to 40
km during the electrification. The resulting upward flux of small positive ions in this region
exhibits a positive divergence between 20 and 40 km altitude due to the rapid increase in ion
mobility with altitude that accompanies the rapidly decreasing neutral density in this region. This
decreases the positive ion density in the region. Small negative ions experience a corresponding density increase, and net negative space charge results. Of course, the inverse scenario is obtained during a negative CG discharge and, were it only for the flow of ions, we would expect a positive space charge of similar magnitude to develop in that case. This is not observed in the simulation results. In fact, the great disparity in the magnitude of the space charge that develops in the two cases is the result of the electron response.

Owing to their much greater mobility, electrons are able to relocate over significant distances in response to cloud charge-induced high altitude electric fields, on the time scale of the electrification period. In the case of the positive CG discharge, electrons flow downward into the 20 - 40 km altitude region, where they are rapidly converted to small negative ions by the process of three body attachment. Figure 4.7c shows that an initial increment in the electron density over equilibrium values is depleted on sub-millisecond time scales at altitudes below 40 km. The figure also shows that the resulting increment in small negative ion density is resolved over much longer times (hundreds of milliseconds), because an excess of negative ions can be depleted only by the much slower processes of recombination with positive ions. The relaxation time of the system to single-species population increments grows longer with increasing altitude, as the neutral molecules, which serve as reactants or buffers in most of the reactions, become scarcer. At low altitudes, the charge carried by the influx of electrons is quickly transferred to negative ions, which can persist for hundreds of milliseconds.

The rapid transfer of charge from the electrons to negative ions explains the slow response of charge density to the actual discharge, as illustrated in Figure 4.12e. Despite the electric field reversal and growth that takes place during the first milliseconds after the onset of the discharge, the negative space charge dissipates only after several hundreds of milliseconds. This is because the electrons that could swiftly respond to field changes have been converted into relatively immobile ions by the time of discharge onset. This further explains why the influx of electrons does not enhance low altitude conductivity during the positive CG discharge, as illustrated in Figure 4.12c. The principal cause of diminished conductivity in the case of both types of discharge is elevated electron temperature.

In great contrast to the scenario outlined for positive CG discharges, electrons play virtually no role in the medium's response to negative CG discharges. Electrons are present only in extremely low concentrations between 20 and 40 km in the quiescent mesosphere (1.0E+06 to 1.0E+07 m⁻³), but their concentration increases rapidly through this region and above it. Thus,
while a large reservoir of electrons at higher altitudes is available to flow into the region during a positive discharge, only a small number of electrons is available to be expelled during a negative discharge. This is the core of the observed asymmetry between positive and negative CG discharges.

Finally, we note from Figure 4.11 that one result of high altitude charge separation is the elevation of off-axis electron temperatures that fall rapidly with increasing radius for the case of the positive CG discharge. The presence of even small amounts of on-axis space charge at high altitudes, as a result of non-local particle flow, acts as an electric field source that does not otherwise exist. As a result, radial electric fields and electron temperatures are increased in the surrounding region.

4.4.2 Persistent Effects of a Lightning Discharge

Some effects of a lightning discharge can persist for many seconds in the 20 - 70 km altitude region. This is illustrated by the seconds-long decline of the electron temperature at 68.5 km altitude shown in Figure 4.13a. The system stimulus in this case consists of a much longer electrification than we have considered up to this point, but is followed by a standard, positive CG discharge of 5 ms duration. During the electrification period, cloud charge grows parabolically in time for a total of 20 s prior to the onset of lightning. This is a typical occurrence, according to Israel [1973, p. 523]. During the latter part of this electrification, the high altitude electric fields and temperatures have achieved a rather long-term steady state. The discharge is exponential, with an e-folding time of 1.2 ms [Uman, 1987, p. 199]. During the first few milliseconds of the discharge, the high altitude electron temperature increases abruptly. This rapid temperature increase contrasts sharply with the much more gradual decline.

Nevertheless, by the endpoint of the temperature evolution shown in Figure 4.13a (9 s after discharge onset), high altitude temperature has fallen to a few percent of its peak value, and low altitude temperature has returned to its quiescent value (as seen in curve 2 of Figure 4.13c). Charge density and electric fields have also returned to their fair weather profiles by this time. However, conductivity in the 20 - 40 km altitude range remains elevated by as much as two orders of magnitude. This is demonstrated by Figure 4.13b, which also indicates that the momentary trend in this altitude range is toward increasing conductivity with time. Figure 4.13b shows the return of conductivity to its quiescent values at altitudes above 40 km, but a growing departure
Figure 4.13: Persistent effects after positive CG lightning discharge. 200 C lowered to ground from 15 km charge center after 20 s parabolic electrification. (a) Temporal evolution of electron temperature at 68.5 km altitude on simulation axis. Altitude profiles along simulation axis of (b) total conductivity, (c) electron temperature, and (d) the difference in charged particle density from equilibrium values. Numbered curves in (b) and (c) represent conditions at (0) equilibrium, (1) 5.62 s after onset of discharge, and (2) 8.96 s after onset of discharge. (d) Off-plot excursions of \( n_m \) (to the left) and \( n_e \) (to the right) are \( \pm 2.0 \times 10^9 \) m\(^{-3}\).
from quiescent conductivity at lower altitudes. The general trend toward increasing conductivity is the result of the continuing decline of electron temperatures toward their quiescent values, as shown in Figure 4.13c. The particular enhancement of conductivity in the 20 - 40 km altitude range is the result of the dramatic, long-term elevation of electron density in that region, shown in Figure 4.13d.

The large excursions of cluster ion and small negative ion density (± 2.0E+09 m⁻³) about 15 km altitude that appear in Figure 4.13d represent the tiny, but finite, residual positive cloud charge. This effect is confined to altitudes below 20 km. Between 20 km and 40 km altitude, the electron and small ion densities exhibit equal and opposite departures (± 1.0E+08 m⁻³) from their equilibrium values. Charge density in this region is very small (< 1.0E+06 electronic charges per m³) and the electric field has fully relaxed. While the maximum depletion of the small negative ions amounts to only 1% of their equilibrium density, the electron density is increased by one to two orders of magnitude. This accounts for the elevated conductivity in the region.

The origin of the electron density enhancement is the period immediately after the onset of the discharge, during which a strong negative electric field develops above the thundercloud. All of the charged species respond to this field but the electrons are most readily swept away. The few electrons present are quickly depleted, as are the electrons that are subsequently detached from negative ions. As a result of the rapid electron response, the positive ion populations between 20 and 40 km altitude are relatively unaffected. The negative ions, on the other hand, experience significant depletion. By the time of Figure 4.13d, the negative electric fields in the region have disappeared, and electrons have returned to reestablish charge neutrality. The process of three body attachment mutually diminishes both the electron enhancement and the negative ion depletion. Eventually, both populations are restored to their equilibrium values.

The persistent changes in the high altitude medium shown in Figure 4.13 amount to preconditioning the system prior to the next lightning discharge. Subsequent lightning flashes within tens of seconds of the initial flash encounter a medium that may depart dramatically from the quiescent state. The enhanced conductivity in the 20 to 40 km altitude range generally results stronger electron heating in that region, but reduced electric fields and electron temperatures at higher altitudes.
4.4.3 Downward Moving Temperature Depression

We will now examine the first of two dynamic features exhibited by the simulation results. Downward moving features have been observed in high time resolution, photometric measurements of sprite and elves events [Fukunishi et al., 1996]. Unfortunately, these observations are not directly comparable to the simulation results presented here. Time correlated, in-situ measurements over a significant spatial baseline would be required to validate in detail the physical mechanisms proposed here. It is nevertheless true that the observed feature speeds are consistent with the simulation predictions.

The first of the two features is a narrow arc of depressed electron temperature that appears to move rapidly downward from high altitudes toward the cloud source during the discharge period. Figure 4.14a shows a snapshot of electron temperature at all locations in the simulation volume taken 0.5 ms after the onset of a lightning discharge that lowered 50 C of positive charge to ground at a peak current of 40 k A. An arc of depressed temperature is observed in this figure, beginning on the simulation axis at an altitude of 63 km and moving slightly downward with increasing radius. The minimum temperature is 25.2 meV, compared to the peak high altitude temperature of 441 meV and a local initial temperature of 21.2 meV. The downward motion of the temperature depression at the simulation axis is charted in Figure 4.14c. At the highest altitudes, the speed of the feature along the axis is 5.3E+07 m/s. Such features are observed for both positive and negative CG discharges.

Examination of corresponding results for the electric field shows that the temperature depressions coincide with nulls in the electric field, which move toward the location of the cloud charge during the discharge period. The immediate response of the electron temperature to the electric field stems from our steady state solution of the electron energy transport equation. We should thus expect that a temperature depression will precisely follow an electric field null. Moreover, the field nulls and their motion are the direct result of the electric field reversal that occurs during the discharge period and ultimately affects all altitudes. The rapid withdrawal of positive charge from the cloud during the discharge causes a reduction in the upward and radially outward-directed (positive) electric field that is established by the end of the electrification period. Due to the slowness of the dielectric response of the low altitude plasma (e.g., ~100 ms at 48 km altitude according to Figure 4.7e) relative to the millisecond charge removal time scale, the electric field becomes less positive everywhere above the cloud. This results in field reversal, first at the
Figure 4.14: Simulation results showing downward movement of temperature depression during 40 kA, positive cloud-to-ground discharge. (a) Electron temperature versus position at 0.5 ms after onset of discharge. (b) Vertical electric field (V/m) along simulation axis at 0.5 ms after onset of discharge. Rapid removal of cloud charge causes electric field curve to shift downward, forcing field null to lower altitude. (c) Altitude of temperature depression versus time after discharge onset.
highest altitudes and largest radii, where the field magnitude is smallest. As charge removal continues, the field reverses at locations closer and closer to the cloud itself, resulting in the apparent motion of the field null and the associated temperature depression. This process is illustrated in Figure 4.14b.

For the purpose of validating this interpretation of the moving temperature depressions, we can estimate the speed at which a high altitude electric field null would appear to move as a result of cloud-height charge removal. If we limit consideration to the lower half of the simulation system, well away from the upper conductive boundary and the cloud charge itself, we can model the change in the local electric field as equal to the change in the field of a vertical electric dipole as the dipole moment is reduced by charge removal. This is true in spite of the fact that the total existing field is not that of a dipole, due to shielding by the medium. The time rate of change of vertical electric field, \( E_z \), on the system axis due to lightning current \( I \) is obtained from the standard result for a dipole field [Zahn, 1987, p. 58].

\[
\frac{\partial E_z}{\partial t} = \frac{Iaz}{\pi \varepsilon_0 \left( z^2 - a^2 \right)^2}
\]  

(4.5)

where \( a \) is the altitude of the cloud charge, and \( z \) is the instantaneous altitude of the field null. The rate of altitude variation of the vertical electric field is given by the Poisson Equation in cylindrical coordinates as:

\[
\frac{\partial E_z}{\partial z} = \frac{\rho}{\varepsilon_0}
\]  

(4.6)

where \( \rho \) is the local charge density. We combine (4.5) and (4.6) to obtain the apparent speed of the field null on the axis:

\[
\frac{dz}{dt} = \frac{Iaz}{\pi \rho \left( z^2 - a^2 \right)^2}
\]  

(4.7)

For a current, \( I \), of 40 kA, cloud height, \( a \), of 15 km, null altitude, \( z \), of 60 km, and simulated charge density, \( \rho \), of 1.5E-14 C per m³, this yields a speed of 6.7E+07 m/s, in reasonable agreement with the simulation finding. It is found from the simulation that the vertical speed of the temperature depression varies only weakly with peak cloud charge. This is because the discharge current is exactly proportional to the peak cloud charge in the simulation model, and the equilibrium charge density at the end of the electrification period is very nearly so. Thus, changes in peak cloud charge produce proportional changes in \( I \) and \( \rho \), which divide out according to (4.7)
and have no effect on the speed of the field null. However, the simulation also shows that the time constant of the lightning discharge strongly affects the apparent speed of the temperature depression. This is consistent with (4.7), in which the speed of the null is proportional to the discharge current, which is inversely proportional to the time constant of the exponential discharge.

Based upon the results of Green et al. [1996], we expect that the predicted electron temperature depression will translate directly into a modification of the associated optical emission spectrum. We therefore propose that the motion of the temperature depression may cause the observed downward motion of the apparent lower boundary of “elves,” as reported by Fukunishi et al. [1996]. This boundary exhibits downward motion at speeds of 0.1 to 0.3 c, which is consistent with the simulation predictions presented in this section.

4.4.4 Downward Propagating Electron Density Discontinuity

The final phenomenon predicted by the simulation is a downward propagating structure in the electron density. The structure has the appearance of a shock-like discontinuity, with lower density ahead of the transition and higher density in its wake. It is confined to within a 15 km radius of the simulation axis, and moves vertically downward.

Figure 4.15a shows a snapshot of the change in the simulated electron density with respect to the initial density, taken at 279 ms after the onset of a positive CG discharge that lowered 200 C to ground with an e-folding time of 1.25 ms. Visible in the figure is a distinct discontinuity in the electron density that has propagated downward from approximately 60 km altitude to its current location at 40 km altitude. Figure 4.15b shows several profiles of density versus altitude at successive times during the relaxation period. The magnitude of the density discontinuity is seen to grow as it propagates downward with a roughly constant speed of 1.0E+05 m/s.

This feature is observed only in simulations of positive cloud-to-ground discharges that lower more than 50 C of charge. It originates within the first hundred milliseconds after the discharge onset, and reaches its low altitude terminus (typically 30 km) in a 300 ms duration. The propagation speed is approximately constant across cases of varying total cloud charge lowered to ground.
Figure 4.15: Simulation results showing downward-propagating discontinuity in electron density during system relaxation following a 160 kA peak current positive cloud-to-ground discharge. (a) Change in electron density relative to initial values versus position at 279 ms after onset of discharge. Discontinuity is at 48 km altitude. (b) Successive altitude profiles of change in electron density along the simulation axis at three times during relaxation period. Upper broken curve at 212 ms after discharge onset, solid curve at 279 ms (corresponding to (a)); lower broken curve at 346 ms.
A downward propagating discontinuity in the electron density may be explained by a nonlinear mechanism akin to an ionization wave. In the present case, a retreating region of dissociative attachment maintains the low electron densities ahead of the transition. A one dimensional, nonlinear attachment wave model is developed in Chapter 5. Such a model can provide a detailed explanation for the phenomenon presented here.
Chapter 5

One-dimensional Model of an Attachment Wave

In Section 4.4.4 we found that the three-dimensional simulation predicts a downward propagating discontinuity in the electron density that follows the simulation axis. Sufficiently close to the axis of our cylindrical simulation volume, the effects of curvature and nonuniformity in the radial dimension disappear, and all significant spatial variation is in the vertical dimension. We will focus on this limit and seek a one-dimensional model to explain the propagating density feature.

Our goal in this undertaking is not to account for an exhaustive catalog of physical effects as was done for the simulation. Instead, we make additional, justifiable simplifications in order to concentrate our attention on the essential processes, in hopes of obtaining an analytical solution.

We will find that in the vicinity of the downward-propagating, shock-like structure, a balance is struck between advection current and drift current at the location of a sharp, propagating gradient in the dissociative attachment. In essence, dissociative attachment is extinguished in the wake of this wave, leaving behind an equilibrium state with higher electron density.

We begin by briefly examining previous approaches to the related problem of propagating ionization wave fronts. We then turn to the new model.

5.1 Previous Approaches to Ionizing Wave Studies

Turcotte and Ong [1968] obtain stationary solutions for the structure of an ionizing wave front. Albright and Tidman [1972] consider a special subclass of the steady-state traveling waves found by Turcotte and Ong [1968], namely, the asymptotic waves to which the system evolves if initially disturbed. The underlying models employed in these two efforts are similar, and both assume constant electron temperature throughout the system. This assumption is utterly invalid under the circumstances considered here, and our principal contribution will be to incorporate the effects of variable electron temperature.
Turcotte and Ong [1968] consider a medium composed of only two components, electrons and positive ions. They use a two-fluid formulation in which the ions are stationary, and the electron dynamics are completely characterized by steady state drifts and diffusion. The electron and ion densities, \( n_e \) and \( n_i \), respectively, are governed by continuity equations, each of which includes a net impact ionization source of the form:

\[
K_p n_e N - K_R n_i^2 n_i
\]

(5.1)

where \( K_p \) and \( K_R \) are production and recombination rate coefficients, respectively, and the first term represents ionization while the second term represents recombination. Since the electron temperature is assumed fixed in this treatment, \( K_p \) and \( K_R \) are constants. This aspect of the approach is so unrealistic that it prohibits comparison of their results to experiment or the to the simulation results presented here. We note that Turcotte and Ong [1968] require recombination to be a three-body process. This is inappropriate for our system, since all of the recombination processes that apply in the mesosphere and lower ionosphere are two-body in nature (see Section 2.6.1).

The system is closed by the Poisson equation, which determines the longitudinal electric field, \( E \), in terms of the difference between the ion and electron densities. In this manner, Turcotte and Ong [1968] obtain three equations for the three unknowns, \( n_e \), \( n_i \), and \( E \). Since they specify their system to be isothermal, the prescription \( T_e = \text{constant} \) replaces an energy equation. They solve their set of governing equations in the temporal steady state. The propagation speed of the resulting wave is related to an eigenvalue of the solution of the nonlinear equations.

Albright and Tidman [1972] focus on the evolution of the system from an initial state to the asymptotic, propagating steady state found by Turcotte and Ong [1968]. They likewise consider a two-component, fluid medium governed by continuity and the Poisson Equation. Moreover, they use the net ionization source of (5.1), including three-body recombination. However, since they are concerned with the temporal development of the traveling wave, they obtain electron velocity from the time-dependent electron momentum equation—not the steady state form.

Albright and Tidman [1972] also incorporate a fluid electron energy transport equation similar to (2.154) in their model. The only energy loss mechanism for electrons is ionization. They do not use the energy transport equation to determine the electron temperature, however, since temperature is assumed to be fixed. Instead, they use it in place of the electron continuity equation to determine the electron density as a function of the other fluid variables. They justify
this approach by postulating that electron temperature is elevated to a range for which impact
ionization is significant. Since the impact ionization rate grows exponentially with temperature, it
provides extremely strong negative feedback to variations in the electron temperature. In effect, a
small increase in electron temperature causes a great increase in the electron energy loss, driving
the temperature back to its initial value. They therefore take the temperature as fixed and use the
steady state electron energy equation to determine the electron density.

This approach is valid only if the electron temperature is set at more than several electron
volts, so that ionization loss becomes significant. In fact, in the Albright and Tidman [1972]
approach, electron temperature must always have been at an extremely elevated level—at least by
mesospheric standards. We know that this is not the case in our system based upon our simulation
results.

We have seen that while impact ionization does not play a significant role in the effects of
lightning on the lower ionosphere, dissociative attachment and other, lower threshold inelastic
processes are indeed important. In fact, we have seen (Section 4.3.2) that these processes do
provide a feedback mechanism that tends to stabilize temperature, as evidenced by the saturation in
the maximum electron temperature obtained for increasing lightning charge lowered to ground.
However, the system never achieves a state in which temperature is uniformly high and fixed. As a
result, we are forced to consider electron temperature to be variable.

Given the considerable degree to which the previous work [Turcotte and Ong, 1968;
Albright and Tidman, 1972] does not apply to the circumstances of interest here, our extension of
it will be far-reaching.

(1) In the mesosphere, at the electron temperature range of interest, dissociative
attachment is an extremely significant electron energy loss mechanism, as well as an important
electron sink. We have seen that it is responsible for high altitude electron depletions (Section
4.1.1). We therefore will include it in our one dimensional model.

(2) Since the dissociative attachment of electrons to oxygen transfers particles from the
electron to the negative ion population, we must include negative ions as well as positive ions and
electrons in our model, making it a three-fluid approach.

(3) Because electron temperature is a critical variable, we include it—as a variable. An
electron energy equation is used to govern the electron temperature.

176
(4) Particle sources due to ionization, attachment, and recombination will be modeled with a temperature dependence that is physically accurate. This will permit quantitative comparison to the simulation results. Moreover, two-body recombination will be included instead of three-body.

5.2 Attachment Wave Model

Motivated in our approach by the discussion of the preceding section, we now present our one-dimensional attachment wave model. We will borrow heavily from our multiple fluid development of Section 2.7, only now we limit our consideration to a single spatial variable, z. We include three fluid species, electrons and positive and negative ions. The dynamics of each species is in the drift-diffusion limit so that we can write the flux of particles of species $\alpha$ as:

$$\Gamma_\alpha = -\frac{\tau_\alpha}{m_\alpha} \frac{\partial}{\partial z} \left( n_\alpha kT_\alpha \right) + \frac{q_\alpha n_\alpha \tau_\alpha}{m_\alpha} E_z$$

(5.2)

$\Gamma$ is flux, $\tau$ is mean time between collisions with neutrals, $m$ is particle mass, $n$ is particle density, $T$ is temperature, and $q$ is charge. The analytical system is further governed by equations of continuity for each species:

$$\frac{\partial n_e}{\partial t} = -\frac{\partial}{\partial z} \Gamma_e + G' - D$$

(5.3)

$$\frac{\partial n_p}{\partial t} = -\frac{\partial}{\partial z} \Gamma_p + G'$$

(5.4)

$$\frac{\partial n_m}{\partial t} = -\frac{\partial}{\partial z} \Gamma_m + D$$

(5.5)

where subscripts $e$, $p$, and $m$ refer to electrons, positive ions, and negative ions, respectively, and $G'$ is the volume rate of impact ionization of neutrals, minus electron-positive ion recombination. $D$ is the rate of dissociative attachment. The z-directed electric field is determined by the Poisson Equation.

$$\frac{\partial E_z}{\partial z} = \frac{e}{\varepsilon_0} \left( n_p - n_e - n_m \right)$$

(5.6)

At the mesospheric altitudes in which the simulation predicts a downward propagating electron discontinuity; namely 30 - 60 km, electrons are a minor constituent. That is, $n_e \ll n_p, n_m$, and the densities of the positive and negative ions vary slowly in comparison to the variation of
their difference. The idealized altitude variation of the electron and ion densities is illustrated in Figure 5.1a. This leads us to focus on the difference

$$n_i = n_e - n_m$$  \hspace{1cm} (5.7)

Since the temperature and density of one ion species differ only slightly from those of the other, it is true that:

$$\frac{\partial}{\partial z} \left( n_e T_e - n_m T_m \right) \ll \frac{\sigma_i}{\kappa_c} \left| E_z \right|,$$

where

$$\sigma_i = \frac{e n_e \tau_e}{m_e} + \frac{e n_m \tau_m}{m_m}$$  \hspace{1cm} (5.8)

where \( \tau \) and \( M \) are generic ion collision time and mass, respectively. Using a conductivity (\( \sim 1.0 \times 10^{-08} \) per \( \Omega \cdot m \)) and an ion-neutral collision time (20 \( \mu s \)) characteristic of 60 km altitude, along with an electric field of 1 V/m, the right hand side of the inequality evaluates to \( \sim 1.0 \times 10^{13} \) K per m\(^4\), which is large compared to the left hand side of \( \sim 5.0 \times 10^{11} \) K per m\(^4\) for a typical ion density of \( 1.0 \times 10^{09} \) per m\(^3\), a temperature of 0.5 eV, and a scale length of 10 m.

Using the electron flux (5.2) in (5.3) and combining (5.4) and (5.5) under the condition of (5.8) yields a three dimensional system in \( n_e, n_i, E_z \), and \( T_e \).

$$\frac{\partial n_e}{\partial t} = \frac{\partial}{\partial z} \left( \frac{\tau_e}{m_e} \frac{\partial}{\partial z} \left( n_e T_e \right) + \frac{e n_e \tau_e}{m_e} E_z \right) + G' - D$$  \hspace{1cm} (5.9)

$$\frac{\partial n_i}{\partial t} = -\frac{\partial}{\partial z} \left( \frac{\sigma_i}{e} E_z \right) + G' - D$$  \hspace{1cm} (5.10)

$$\frac{\partial E_z}{\partial t} = -\frac{e}{\varepsilon_0} (n_i - n_e)$$  \hspace{1cm} (5.11)

In contrast to Turcotte and Ong [1968], electron temperature in this formulation is a dependent variable. We must close this set of equations by using the energy transport equation.

We are able to eliminate the combination of terms, \( G' - D \), from the electron continuity equation (5.9) by combining it with the ion continuity equation (5.10). We thereby obtain:

$$\frac{\partial}{\partial t} \left( n_e - n_i \right) = \frac{\partial}{\partial z} \left( \frac{\tau_e}{m_e} \frac{\partial}{\partial z} \left( n_e T_e \right) + \frac{e n_e \tau_e}{m_e} E_z + \frac{\sigma_i}{e} E_z \right)$$  \hspace{1cm} (5.12)

We now impose the restriction that we are looking for only stationary wave solutions, that is, solutions expressible in the form \( f = f(\xi) \), where \( \xi = z - Vt \), and \( V \) is the propagation speed of the wave in the positive \( z \) direction. As a result of this transformation, speed continues to be
Figure 5.1: Elements of the one-dimensional attachment wave model. (a) Variation of electron and negative and positive ion density, \(n_e, n_m,\) and \(n_p,\) respectively, in altitude region of wave solution. (b) Approximation to electron energy equation. Inelastic processes lead to saturation of electron temperature at high values of electric field. The system under consideration is in the elastic limit.
measured in the earth frame, while position is measured in a frame that moves with the wave. The resulting system of equations may be written:

\[-V \frac{d}{d \xi} (n_e - n_i) = \frac{d}{d \xi} \left( \frac{\kappa \tau_e}{m_e} \frac{d}{d \xi} (n_e T_e) + \frac{e \tau_e}{m_e} n_e E_z + \frac{\sigma_i}{e} E_z \right) \]  

(5.13)

\[-V \frac{dn_i}{d \xi} = -\frac{d}{d \xi} \left( \frac{\sigma_i}{e} E_z \right) + G' - D \]  

(5.14)

\[
\frac{d E_z}{d \xi} = \frac{e}{e_0} (n_i - n_e) \]  

(5.15)

Although we have not yet formally introduced the electron energy equation in the present context, we can proceed with these three equations, bearing in mind that \( T_e, \tau_e, \sigma_i, \) and \( G' - D \) are all functions of the independent variable \( \xi \). We can integrate (5.13) directly to obtain:

\[-V (n_e - n_i) = \left( \frac{\kappa \tau_e}{m_e} \frac{d}{d \xi} (n_e T_e) + \frac{e \tau_e}{m_e} n_e E_z + \frac{\sigma_i}{e} E_z \right) + \text{constant} \]  

(5.16)

We will evaluate the integration constant at the asymptotic high altitude state, \( \xi \to +\infty \). Naturally, this is not a real location in the physical system. It represents conditions far to the high altitude side of the propagating wave. In order to evaluate the terms in (5.16), we must first establish the values of the dependent variables in the high altitude limit. We demand that for any physical solution, the derivative with respect to \( \xi \) must vanish as \( \xi \to +\infty \). Furthermore, we are guided by our knowledge of the real ionosphere to demand that conditions at sufficiently high altitude must be quasi-neutral and field free. Electron temperature will approach some finite value in the high altitude limit.

Thus, in the high altitude asymptotic state we have:

\[
\frac{d n_i}{d \xi} = 0; \quad n_e = n_i = n_0; \quad E_z = 0; \quad \text{and} \quad T_e = T_0 \]  

(5.17)

We note that this state is actually an equilibrium of the system described by (5.13), (5.14), and (5.15). Since we are looking for shock solutions, it is natural for us to associate an equilibrium condition with each of the asymptotic states of the system, \( \xi \to \pm \infty \). It is the nature of a shock solution that far from the vicinity of the shock transition, the system approaches a flat, equilibrium state.

Turning now to the low altitude asymptotic state of the system, we propose that appropriate values of the dependent variables there may be obtained by extrapolating the physical trends at 30 km downward indefinitely. In this manner, we conclude that electron density should
approach zero as $\xi \to -\infty$. The same is true of the difference between the positive and negative ion densities, $n_i$. Quasi-neutrality is again obtained in the low altitude state. Both the electric field and the electron temperature approach finite values. These low altitude asymptotic conditions are listed together with the high altitude ones in Table 5.1:

<table>
<thead>
<tr>
<th>LOW ALTITUDE</th>
<th>HIGH ALTITUDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi \to -\infty$</td>
<td>$\xi \to +\infty$</td>
</tr>
<tr>
<td>$n_e = 0$</td>
<td>$n_e = n_0$</td>
</tr>
<tr>
<td>$n_i = 0$</td>
<td>$n_i = n_0$</td>
</tr>
<tr>
<td>$E_z = E_0$</td>
<td>$E_z = 0$</td>
</tr>
<tr>
<td>$T_e = T_1$</td>
<td>$T_e = T_0$</td>
</tr>
</tbody>
</table>

Table 5.1: Asymptotic states of the one-dimensional system.

The table suggests that the electron density, as well as the difference in the ion densities, goes from a finite value at high altitudes to zero at low altitudes. The electric field, on the other hand, goes from the completely relaxed value of zero at high altitudes to some finite value at low altitudes. The electron temperature takes on values determined by the other variables in the asymptotic states. Table 5.1 captures the behavior of the electric field and the particle densities at large distances from the wave. To be specific, $E_0$ is interpreted as the known value of electric field impinging on the mesosphere from below, subsequent to a lightning flash. Similarly, $n_0$ is taken to be the electron density in the vicinity of the 60 km altitude at which the ionization front is observed to originate in the simulation.

We use the high altitude asymptotic state to evaluate the integration constant in (5.16), as promised. The constant is found to be zero and we rearrange (5.16) to obtain:

$$\frac{d}{d\xi} (n_e \kappa T_e) = \frac{V m_i}{\tau_e} (n_i - n_e) - e n_e E_z - \frac{m_i \sigma_i}{e \tau_e} E_z$$  \hspace{1cm} (5.17)

Now, if we assume that the ion conductivity is uniform, we can use the Poisson Equation (5.15) to rewrite the ion continuity equation (5.14) as

$$\frac{d n_i}{d\xi} = \frac{\sigma_i}{\varepsilon_0 V} (n_i - n_e) - \frac{G^i - D}{V}$$ \hspace{1cm} (5.18)

The assumption of uniform ion conductivity is a major simplification. However, it does not interfere with our investigation of the dependent variables since ion conductivity varies neither with electron temperature, nor with the difference in the ion densities. Instead, it depends on the total ion density, which is constant in comparison to $n_i$, and on the neutral density. Uniform ion conductivity is thus equivalent to locally uniform neutral density. In fact, we will go one step
further in the present analysis and assume that the ion drifts represented by the first term in (5.18) are insignificant in comparison to the production term. This is consistent with our assessment in Section 2.7.5, and will serve to demonstrate that ion drifts are not important to the propagation of an electron shock. Thus, we reduce (5.18) to

\[
\frac{d n_e}{d \xi} = -\frac{G' - D}{V} \quad (5.19)
\]

By the same token, assuming stationary ions eliminates the ion drift term from (5.17).

It is now necessary to specify the net impact ionization and dissociative attachment terms in (5.19). In creating the new dependent variable \( n_t \) to represent the difference between the densities of the positive and negative ions, we were able to isolate all of the particle sources and sinks in the combination of terms \( G' - D \). This quantity represents the difference between the rate of impact ionization, and the total rate of dissociative attachment and recombination. In the language of our chemical rate equations (2.110) it would be written:

\[
G' - D = G - \alpha_3 N_{O2} n_e - \alpha_D N_{O2} n_e - \alpha_n n_p n_e \quad (5.20)
\]

where \( \alpha_3 \) is the three-body attachment rate, \( \alpha_D \) is the dissociative attachment rate, and \( \alpha_n \) is the electron-ion recombination rate. We have ignored those processes that were found to be less significant in Section 2.6.2. The processes represented in (5.20) are those of primary importance. The impact ionization rate \( G \) is itself proportional to the electron and the (constant) neutral densities according to (2.101).

\[
G = R N n_e \quad (5.21)
\]

so

\[
G' - D = \left[ R N - \alpha_3 N_{O2} \right]^2 n_e - \alpha_n n_p n_e \quad (5.22)
\]

We can simplify the appearance of this by writing

\[
G' - D = K_1 N n_e - K_2 n_p n_e \quad (5.23)
\]

where \( N \) is neutral density and \( K_1 \) and \( K_2 \) are temperature-dependent rate coefficients. Evaluating this in the low altitude asymptotic limit of Table 5.1, we see that the net particle source rate is zero in the low altitude equilibrium. We expect the same to be true in the high altitude limit. \( K_1 \) may be expressed in terms of \( K_2 \) if we demand that \( G' - D \) be zero at the high altitude equilibrium state:

\[
K_1 N = K_2 \left( n_o + N_m \right) \quad (5.24)
\]

where we have replaced \( n_p \) in (5.23) using (5.7). If we insist that (5.24) is true for arbitrary values of high altitude equilibrium temperature, then it is true for all \( T_e \). Using this result for \( K_1 \), we find
\[ G' - D = K_2 n_e (n_0 - n_i) \]

The temperature dependence of \( K_2 \) is calculated using the collisional effects models developed in Section 2.5. It is constructed to reflect principally the processes of impact ionization and dissociative attachment. In fact, we actually calculate \( K_1 \) and determine \( K_2 \) from (5.24). The end result is that

\[ K_2(T_e) \approx \frac{1}{n_m} \left[ I - \alpha_D(T_e) N_{O2} \right] \]

(5.26)

where \( I \) represents all of the electron sources that maintain equilibrium against high rates of dissociative attachment at low altitudes. Thus, \( K_2 \) represents a net production rate in regions of reduced electron temperature, hence reduced dissociative attachment. \( K_2 \) is always positive, taking on large values where temperature is low, and small values where temperature is high.

Inserting (5.25) in (5.19), we rewrite the three ordinary differential equations governing the system.

\[
\frac{d}{d\xi} \left( n_e T_e \right) = \frac{V m_e}{\tau_e} (n_i - n_e) - e n_e E_z
\]

(5.27a)

\[
\frac{dn_i}{d\xi} = - \frac{K_2 n_e}{V} (n_0 - n_i)
\]

(5.27b)

\[
\frac{dE_z}{d\xi} = \frac{e}{\varepsilon_0} (n_i - n_e)
\]

(5.27c)

where we have eliminated the ion drift term from (5.17) to obtain (5.27a) under the assumption of stationary ions discussed above. We close the above system of equations using the steady state form of the electron energy transport equation (2.154).

\[
0 = -e \gamma E_z - \frac{\delta}{\tau_e} n_e \kappa (T_e - T) - n_e \Phi(T_e)
\]

(5.28)

We have left out the thermal gradient and impact ionization source terms found in (2.154) because they are dwarfed by the elastic collisional loss at the altitudes and temperatures of interest.

The inelastic energy loss term grows extremely rapidly with temperature and sets the maximum temperature at high values of electric field. This is the temperature saturation mechanism we mentioned in Section 5.1. The assumption of saturation is used by Albright and Tidman [1972] and by Turcotte and Ong [1968] to assert constant temperature. In our case, the saturated state is not attained and temperature varies through the wave, but we nevertheless find it useful to distinguish between the saturated state, in which inelastic loss is dominant, and the non-
saturated state, in which elastic loss dominates. We are thereby able to ignore the detailed temperature dependence of Φ by using only the elastic loss term at low values of electric field, and using a constant maximum temperature at large values of electric field. This model is illustrated in Figure 5.1b. The electron temperature that demarcates the two regions is several electron volts, well above the temperatures considered here.

In the low temperature limit, we rewrite (5.28) as:

\[
E_z = \frac{m_e \delta}{e^2 \tau_e} \kappa (T_e - T)
\]  

(5.29)

where we have used (5.2) to replace the electron flux, ignoring the pressure gradient term relative to the kinetic energy term in accordance with the discussion of Section 2.7.5. We use a simplified model for the temperature dependence of the electron momentum transfer collision rate, based upon a mean value \( \sigma_c \) for the electron-neutral momentum transfer collision cross section [Krall and Trivelpiece, 1986, p. 321].

\[
\tau_e = \frac{1}{N \sigma_c} \sqrt{\frac{m_e}{\kappa T_e}} = \tau_{e0} \sqrt{\frac{T}{T_e}}
\]

(5.30)

where \( \sigma_c = 4.0 \times 10^{-19} \text{ m}^2 \), and we have lumped several constants into \( \tau_{e0} \), which is simply \( \tau_e \) for the case \( T_e = T \). Using this model in (5.29) and solving for the electron temperature yields:

\[
\frac{T_e}{T} = \frac{1}{2} + \sqrt{\frac{1}{4} + \gamma^2 \left( \frac{E_z}{E_0} \right)^2} \quad \text{where} \quad \gamma^2 = \frac{e^2 \tau_{e0} E_0^2}{\kappa T m_e \delta}
\]

(5.31)

According to (5.31), the electron temperature \( T_e \) at the high altitude asymptotic state, where \( E_z = 0 \), is just the background neutral temperature, \( T \):

\[
T_e = T
\]

(5.32)

Similarly, at the low altitude asymptotic state, where \( E_z = E_0 \) by definition, we find that the electron temperature is

\[
T_e = T \left( \frac{1}{2} + \sqrt{\frac{1}{4} + \gamma^2} \right)
\]

(5.33)

Moreover, for the conditions at 60 km altitude, namely \( N = 6.7 \times 10^5 \text{ m}^{-3} \), \( \delta = 3.6 \times 10^{-5} \) for NO\(^+\), and an electric field of 1 V/m, we find that \( \gamma \sim 10 \), so

\[
T_e \approx \gamma T
\]

(5.34)

By the same token, we approximate (5.31) near the low altitude limit as:
\[
\frac{T_e}{T} \approx \gamma \left\lvert \frac{E_z}{E_0} \right\rvert
\]  

Finally, we normalize the equations by taking \( L \) as the scale for length, \( n_0 \) for densities, the magnitude of \( E_0 \) for electric field, and high altitude electron temperature \( T_0 \) for \( T_e \). The resulting equations appear:

\[
\frac{d}{d\xi} (\tilde{n}_e \tilde{T}_e) = \frac{\varepsilon}{\beta} (\tilde{n}_i - \tilde{n}_e) - \frac{1}{\beta} \tilde{n}_e \tilde{E} \]  

(5.36a)

\[
\frac{d\tilde{n}_i}{d\xi} = \bar{g} \tilde{n}_e (\tilde{n}_i - 1) \]  

(5.36b)

\[
\frac{d\tilde{E}}{d\xi} = \frac{1}{\alpha} (\tilde{n}_i - \tilde{n}_e) \]  

(5.36c)

\[
\tilde{T}_e = \sqrt{2 + \sqrt{1 + \gamma^2 \tilde{E}^2}}
\]  

(5.37)

where the tilde indicates a normalized variable, and the dimensionless parameters are defined by:

\[
\alpha = \frac{\varepsilon_0 E_0}{en_0 L}, \quad \beta = \frac{\kappa T_0}{e|E_0|L}, \quad \varepsilon = \frac{\sqrt{\gamma m_e c}}{\tau_e |E_0|}, \quad \text{and} \quad \bar{g}(T_e) = \frac{K_2 n_0 L}{V}
\]  

(5.38)

where \( \varepsilon \) and \( \bar{g} \) carry the sign of the propagation speed \( V \). The parameters identified as \( \alpha \) and \( \beta \) are equivalent to the like-named quantities in Section 2.7.4. We obtain the value for \( \alpha \) given in (5.38) if we use the electron dielectric relaxation time as the time scale in the version of \( \alpha \) from Section 2.7.4. Thus, \( \alpha \) measures drift distance versus feature size and \( \beta \) represents particle thermal energy relative to electrostatic potential energy, as before. The new parameter, \( \varepsilon \), measures propagation speed relative to drift speed and depends on the as yet undetermined wave propagation speed, \( V \).

Using (5.37) to determine \( \tilde{T}_e \) in terms of \( \tilde{E} \), equations (5.36) become a three dimensional system of ordinary differential equations. We note that the system has two fixed points, given by

high altitude fixed point: \( (\tilde{n}_e, \tilde{n}_i, \tilde{E}) = (1,1,0) \)  

(5.39)

low altitude fixed point: \( (\tilde{n}_e, \tilde{n}_i, \tilde{E}) = (0,0,\pm 1) \)  

(5.40)

where the sign of the electric field at the low altitude fixed point is the same as that of the impinging low altitude electric field, so that there is only one low altitude fixed point—it may simply be in one of two locations. As anticipated, there is a fixed point corresponding to each of the asymptotic conditions listed in Table 5.1. Since the asymptotic states are the solution values at
\[ \xi \rightarrow \pm \infty, \text{ and since a physical solution must exist for all } \xi, \text{ it must be the case that any physical trajectory for the system will include the fixed points (5.39) and (5.40). Thus, to discover the asymptotic behavior of the solution, we must examine (5.36) near the fixed points. To that end, we consider the coefficients to be constant in the vicinity of the fixed points and linearize in their neighborhood.} \]

We linearize the system (5.36) near the high altitude fixed point using the following substitutions:

\[ \tilde{n}_e = 1 + \Delta n_e \]  
\[ \tilde{n}_i = 1 + \Delta n_i \]  
\[ \tilde{E} = \Delta E \]  
\[ \tilde{T}_e = 1 + \gamma^2 (\Delta E)^2 \]

where the electron temperature has no linear term near the fixed point. Inserting these forms for the dependent variables, (5.36) become

\[ \frac{d}{d\xi} \begin{pmatrix} \Delta n_e \\ \Delta n_i \\ \Delta E \end{pmatrix} = \begin{pmatrix} -\varepsilon/\beta & \varepsilon/\beta & -1/\beta \\ 0 & \tilde{g} & 0 \\ -1/\alpha & 1/\alpha & 0 \end{pmatrix} \begin{pmatrix} \Delta n_e \\ \Delta n_i \\ \Delta E \end{pmatrix} \]  

(5.42)

The eigenvalues, \( \lambda \), for this equation are:

\[ \lambda = \begin{pmatrix} \lambda_+ \\ \lambda_- \end{pmatrix} = -\frac{\varepsilon}{2\beta} \pm \sqrt{\left(\frac{\varepsilon}{2\beta}\right)^2 + \frac{1}{\alpha \beta}} \]

(5.43)

For the electron temperature associated with the high altitude fixed point, \( \tilde{g} \) is finite and negative for a wave propagating downward. The remaining roots are always real and of opposite sign, indicating that the high altitude fixed point is always a saddle point. In order for the solution to exist at all \( \xi \), it must arrive at the high altitude fixed point as \( \xi \rightarrow +\infty \). This condition specifies a unique trajectory for the system, namely the one that lies along the eigenvector corresponding to the negative eigenvalue. If the wave is propagating downward, so that \( \varepsilon < 0 \), then \( \lambda_- \) is the negative eigenvalue. The corresponding eigenvector is given by

\[ \begin{pmatrix} \Delta n_e \\ \Delta n_i \\ \Delta E \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -(\varepsilon + \lambda_- \beta) \end{pmatrix} \]

(5.44)
This situation is illustrated in Figure 5.2, which plots system trajectories obtained by Runge-Kutta integration (see Section 3.4) of (5.36) near the upper fixed point. Figure 5.2 shows a three dimensional plot of representative trajectories in the \((\vec{n}_e, \vec{n}_i, \vec{E})\) phase space. The trajectories all bend and miss the upstream saddle point—except for the dashed trajectory, which asymptotically approaches the saddle point along the eigenvector (5.44). The dashed trajectory represents the shock-like solution.

### 5.2.1 Solutions for Positive Low Altitude Electric Field

In order to linearize the system near the low altitude fixed point, we must specify whether the low altitude electric field is positive or negative. Only the magnitude of this field enters in the normalization (5.38). In the case of a positive field, our linearizing substitution appears:

\[
\begin{align*}
\tilde{n}_e &= \Delta n_e \\
\tilde{n}_i &= \Delta n_i \\
\tilde{E} &= 1 + \Delta E \\
\tilde{T}_e &= \gamma (1 + \Delta E) \text{ for } \gamma \gg 1
\end{align*}
\]

The linearized equations then become

\[
\frac{d}{d\xi} \begin{pmatrix} \Delta n_e \\ \Delta n_i \\ \Delta E \end{pmatrix} = \begin{pmatrix}
\frac{1 + \epsilon}{\gamma \beta} & \frac{\epsilon}{\gamma \beta} & 0 \\
-\frac{\epsilon}{\gamma} & 0 & 0 \\
-\frac{1}{\alpha} & \frac{1}{\alpha} & 0
\end{pmatrix} \begin{pmatrix} \Delta n_e \\ \Delta n_i \\ \Delta E \end{pmatrix}
\]

The associated eigenvalues are:

\[
\lambda = \begin{cases}
\lambda_+ = \left(1 + \frac{\epsilon}{2\gamma \beta}\right) \pm \sqrt{\left(1 + \frac{\epsilon}{2\gamma \beta}\right)^2 - \frac{\epsilon \tilde{g}}{\gamma \beta}} \\
\lambda_- = 0
\end{cases}
\]

The second term under the quadratic is always positive for \(K_2 > 0\), so that for \(\tilde{g}\) of sufficient magnitude, the eigenvalues would be complex, leading to oscillatory behavior in the low altitude limit. In practice, \(\tilde{g}\) is small at low altitudes, and this condition is not observed. Instead, the eigenvalues are either both positive or both negative, depending upon the sign of the first term. In the case that they were both negative, the low altitude fixed point would be a sink and the trajectory
Figure 5.2: 3-D representation of \( (\vec{n}_e, \vec{n}_i, \vec{E}) \) phase plane. Asterisks mark low altitude endpoints of calculated trajectories. Dashed trajectory approaches upstream fixed point asymptotically, and corresponds to integrated solution shown in Figure 5.3.
originating there could never leave. This is an unphysical situation, and we must require the first term in \( \lambda_z \) to be positive. This places a condition on \( \varepsilon \).

\[
\varepsilon < -1
\]  

(5.48)

Via (5.38), this in turn places a condition on the propagation speed, which is the only free parameter in the problem.

\[
\nu < -\frac{\tau_e e |E_0|}{m_e}
\]  

(5.49)

In other words, in the presence of a positive low altitude electric field, the wave must propagate at a speed greater than the downward electron drift speed. Only downward propagating solutions are admitted. This is an understandable condition, in that if the wave were propagating upward, or downward more slowly than the drift speed, then electrons would be drawn continuously downward from the shock front, and steady state could not be achieved.

### 5.2.2 Solutions for Negative Low Altitude Electric Field

In the case that the low altitude electric field is negative, so that the associated fixed point is located at \( \tilde{E} = -1 \), the linearized equation may be written:

\[
\frac{d}{d\xi} \begin{pmatrix} \Delta n_e \\ \Delta n_i \\ \Delta E \end{pmatrix} = \begin{pmatrix} \frac{1-\varepsilon}{\gamma \beta} & \varepsilon/\gamma \beta & 0 \\ -\tilde{g} & 0 & 0 \\ -1/\alpha & 1/\alpha & 0 \end{pmatrix} \begin{pmatrix} \Delta n_e \\ \Delta n_i \\ \Delta E \end{pmatrix}
\]  

(5.50)

and the eigenvalues are:

\[
\lambda = \left\{ \begin{array}{c}
\lambda_+ = \frac{1-\varepsilon}{2\gamma \beta} \pm \sqrt{\left(\frac{1-\varepsilon}{2\gamma \beta}\right)^2 - \frac{\varepsilon \tilde{g}}{\gamma \beta}} \\
0
\end{array} \right.
\]  

(5.51)

The condition for a physical solution—that both eigenvalues \( \lambda_\pm \) be positive—implies:

\[
\varepsilon < 1
\]  

(5.52)

For an upward propagating wave, this means that the wave speed cannot exceed the upward drift speed of the electrons. Just as in the case of the positive low altitude electric field, this prevents the drift loss of electrons downward through the wave front. Downward propagating waves are also admitted by this criterion and can have any propagation speed.
By rewriting (5.36a), we can view the steady state achieved by the propagating shock solution in terms of current balance in the wave frame

$$0 = \frac{d}{d\xi} \left( \tilde{n}_e \tilde{T}_e \right) - \frac{e}{\beta} \left( \tilde{n}_i - \tilde{n}_e \right) + \frac{1}{\beta} \tilde{n}_e \tilde{E}$$

(5.53)

The first term represents diffusion current, which is negative for electron diffusion and zero for the cold, motionless ions. The second term represents advection current through the wave, as charge present in the medium flows through the wave opposite the direction of propagation. The third term represents the familiar drift current. For parameters typical of the 30 - 60 km altitude range, we find $\alpha \sim 0.1$, and $\beta \sim 1.0E-04$. For the propagation speeds observed in the simulation, $\varepsilon \sim -1$. Based upon these values of the dimensionless parameters, (5.53) indicates that diffusion does not play a strong role in the wave. This is doubly true since the gradients in density and temperature are directed oppositely. This leaves the advection current to balance the drift current. In the case of negative low altitude electric field, the screening charge density is positive, and flows upward through the downward propagating wave to yield positive advection current. A positive advection current is, indeed, able to counter the downward drift current and maintain a stationary wave.

The exact form of the stationary wave obtained under this scenario is presented in Figure 5.3. It is appropriate to examine the case of negative low altitude electric field, for that is the polarity of the field in the aftermath of the positive cloud-to-ground lightning discharges considered in Chapter 4. Figure 5.3 shows a calculated solution for a wave propagating downward at a speed of $1.9E+05$ m/s in the presence of a negative low altitude electric field of magnitude 1 V/m. The transition region is several hundred meters thick, so that the change in electron density would appear as a sharp discontinuity on the spatial scale of the simulation. The high altitude limit of the calculated solution is not flat due to the necessity for starting at some finite location in $\xi$, and integrating downward along the appropriate eigenvector away from the high altitude saddle point.

The electron and ion densities shown in Figure 5.3a are very nearly equal. This is a result of the fact that $\alpha \ll 1$ so that the characteristic electron drift distance is very small compared to the scale size of the wave. Consequently, gradients in the electric field provide very strong feedback toward reducing the charge density of the system.

We note that the electron temperature in the low altitude limit is relatively high, 1.6 eV. This is a direct result of our ignoring inelastic loss processes in our approximation to the electron energy equation (5.29). The fact that the maximum temperature is below the several eV threshold for intense inelastic loss nevertheless validates the approximation. As anticipated, dissociative
Figure 5.3: Calculated one-dimensional solution profiles for downward propagating attachment wave in presence of -1 V/m low altitude electric field. Electron and net positive ion \((n_p - n_m)\) density (a), longitudinal electric field (b), electron temperature (c), and rate of dissociative attachment (d) are shown versus relative position. Electron density is shown with solid curve in (a), net positive ion density with dashed curve. Net positive ion density is several percent greater than electron density in transition region.
attachment sets in below the altitude of the transition, as demonstrated by Figure 5.3d. Thus, our one-dimensional solution represents a balance between advection and drift at a propagating front of extinguishing dissociative attachment. This conclusion provides some insight into the simulation results of Section 4.4.4.

These results do not directly address issues of the onset and evolution of an initial density disturbance into an asymptotic, downward-propagating stationary wave. Nevertheless, the requirement for positive space charge in the transition region means that the negative space charge remaining immediately after the lightning discharge must be neutralized by shielding charge before the ambient conditions can be consistent with the one-dimensional solution. Thus, the development of a downward propagating stationary wave will be delayed relative to the discharge by roughly the dielectric response time at 50 to 60 km altitude, or 10 to 100 ms. This is consistent with the simulation results.
Chapter 6

Conclusions

In this chapter, we highlight the most important results presented in this work. These results represent major contributions to the theory of quasi-electrostatic processes in the earth-mesosphere-lower ionosphere system. Moreover, we have made significant new applications of that theory to the study of red sprites and related phenomena. Finally, numerical methods were extended and validated for use in the physical system of interest. We will first summarize the contributions and then present recommendations for future research.

6.1 Contributions to Theory

All aspects of the physical modeling effort entailed by this work were undertaken from scratch due to the utter lack of comprehensive, self-consistent, multiple-fluid models applicable over the entire range of altitudes from the ground to 90 km. The effort was greatly complicated by the fact that the system variables, such as the charged particle densities, and the fundamental system parameters, such as the mobilities, vary over many orders of magnitude within the altitude range of interest.

It is by itself a significant contribution to integrate a system of component models in order to self-consistently address the combined effects of particle flow, particle production and loss due to collisional and chemical processes, electron heating, electron energy loss due to inelastic scattering, and electric field creation—all in a medium that accurately resembles the earth-mesosphere-lower ionosphere system. The outcome is a comprehensive, composite model that will find continued application in calculating the properties of a lightning-disturbed medium for use in studies of radio propagation and optical and infrared luminosity.

The one-dimensional, nonlinear attachment wave model developed here, while not comprehensive, captures the essential physics of propagating density disturbances in the aftermath of a positive cloud-to-ground lightning discharge. As such, it provides the basis for a continuing
examination of the conditions under which a lightning disturbed mesosphere can produce stationary waves.

Some of the specific advances that contributed to the development of the comprehensive model may be summarized as follows.

a. The high altitude electric field solution of Section 2.4.3 provides the spatial variation of an electrostatic field in a cylindrically symmetric medium of extremely high, but finite conductivity. The solution stems from an analytical treatment of current continuity in a quasi-neutral region. The medium is extended all the way to infinite altitude, sidestepping the need for image charges to account for the conductive upper boundary. Not only does this provide more accurate results than the corresponding Poisson Equation solution; it is also the physically appropriate method of obtaining the solution, for in the absence of significant high altitude charge, the high altitude electric field is determined by current flow in the presence of varying conductivity.

b. The incorporation of an accurate model of elastic and inelastic energy loss mechanisms in the fluid electron energy transport equation is a novel and sophisticated means of closing the chain of moment equations derived from the Boltzmann equation. In Section 2.5.2 a kinetic theory result for the electron velocity distribution in the presence of an electrostatic field, the Druyvestein distribution, is characterized in terms of its effective temperature. The effective temperature is a well-defined concept of the fluid formulation, and once the collisional energy loss rates are calculated using this distribution, the electron energy transport equation may be solved for the self-consistent electron temperature. This approach is an effective and versatile means of closing the system of fluid equations.

c. A chemical effects model has previously been developed for the ionospheric D region [Rowe et al., 1974, Glukhov et al., 1992]. Because it ignores the variation of atmospheric water vapor content with altitude, it is not useful below 40 km altitude, where it produces wildly erroneous equilibrium densities for the charged particle populations. The chemical effects model of Section 2.6 accurately includes the hydration chain reaction, which is all-important at low altitudes. Because the chemical effects involve time scales much shorter than those of the transport processes, the model must be integrated separately to obtain accurate density increments on the transport time scale.
d. The use of a multiple fluid model to explicitly incorporate such a wide variety of physical mechanisms over such a wide spatial variation in the ambient medium is a novel undertaking. Previous approaches [Pasko et al., 1997] have relied on laboratory measurements of electron swarms to provide electron mobility, ionization, and total excitation directly as a function of applied electric field. This technique is inadequate for our purposes. In the first place, we seek the individual rates for the many excitation processes that withdraw energy from the electron population. Moreover, we wish to know the temperature of the high altitude electrons in the aftermath of a lightning discharge. Only a fluid formulation that explicitly incorporates the collisional processes will provide this information about the disturbed medium.

It is also true that previous approaches [Pasko et al., 1997, and references therein] have only considered in-place changes in particle densities due to collisional processes. Density changes in response to transport were not considered even for highly mobile electrons. The importance of particle relocation is demonstrated repeatedly in the results of Section 4.4. Especially in cases where the evolution of the system over many seconds is of interest, we must include the effects of transport.

The multiple-fluid formulation of Section 2.7 painstakingly includes particle transport, production and loss, as well as heating due to a quasi-static electric field and energy loss to elastic and inelastic processes. A careful examination of conditions in the mesospheric-ionospheric medium leads to a steady state form of the energy transport equation. This algebraic equation for determining the effective electron temperature under the influence of a great many heating and energy loss processes has wide applicability beyond the scope of this work.

e. The one-dimensional attachment wave model of Chapter 5 represents a substantial development beyond the work of Turcotte and Ong [1968] and Albright and Tidman [1972]. Electron temperature is held fixed in their work, but we have found it to be a critical variable. We therefore include it as a variable and use a simplified, low-field energy equation to govern its variation. The results of Section 4.4.4 suggest that the dissociative attachment of electrons to oxygen plays an important role in a propagating electron density discontinuity. Accordingly, we develop a collisional source model that accurately reflects the dominant role of attachment. This, in turn, dictates that we include negative ions as well as positive ions and electrons in our model,
making it a three-fluid approach. Finally, our accurate approach to modeling the collisional electron source permits quantitative comparison to the simulation results.

6.2 Contributions to the Study of Red Sprites and Related Phenomena

Because the observations of sprites have been primarily optical, our simulation results bear upon them only indirectly. However, the simulation results are directly relevant to the electron heating and density changes that are believed to be responsible for sprites. We are able to elucidate these underlying processes, assess their relative importance, and make predictions regarding phenomena yet to be observed.

a. The simulation shows that positive CG lightning discharges lowering 50 C or more of charge to ground are able to produce high altitude electron temperatures exceeding 0.4 eV. According to Green et al. [1996], temperatures between 0.4 and 2.0 eV are consistent with observed sprite optical spectra. Thus, the electron temperatures predicted by the simulation can be expected to produce sprite luminosity. Section 4.2 further shows that the gross morphology of the predicted electron temperature enhancement, coupled with the altitude variation in electron density, is also consistent with sprite observations.

b. The results of Section 4.1.2 demonstrate that the duration of the electric field transient produced by a lightning discharge is a function of the dielectric response of the medium at high altitude, and the chemical response at low altitudes. This has never before been established, because other investigators [Rowland et al., 1996; Pasko et al., 1997] have employed aggregate swarm data instead of an explicit chemical model in their formulation.

c. Peak lightning current has a strong effect on the altitude of the electron temperature peak generated after a positive CG lightning discharge—moving it through a range of 10 km as shown in Section 4.3.3. The effect on the amplitude of the peak is less pronounced. Since peak current is a highly variable lightning parameter, we would expect considerable random variation in the altitude of sprite features. Indeed, all of the altitude dependent sprite features exhibit a measured altitude variance of ± 4 to ± 5 km [Sentman et al., 1995].
d. Section 4.4.1 establishes the importance of high altitude charge separation to the differing effects of positive and negative cloud-to-ground discharges. Typical positive cloud-to-ground lightning discharges can produce electron temperatures at 70 km altitude that are vastly greater than the corresponding temperatures produced by typical negative discharges. We focus on the fact that discharge polarity alone can yield a factor of two difference, and identify the mechanism by which discharge polarity affects electron temperature. That mechanism is charge separation in the 20 to 40 km altitude region. It is clear that charge separation is a critical process underlying the heating of mesospheric electrons, and we are first to point this out.

e. In Section 4.4.2, the multiple-fluid formulation has captured persistent effects of ion and electron redistribution in the mesosphere, subsequent to a lightning discharge. We find long-lasting depletion of negative ions and enhancement of electron populations in the 20 km to 40 km altitude region, accompanied by increases in local conductivity a hundred times above ambient. Remarkably, this magnitude of perturbation persists at least 10 s after the lightning discharge. Persistent effects of this magnitude have not previously been addressed.

f. Finally, the multiple-fluid simulation predicts two new mesospheric phenomena. The first is a band of depressed electron temperatures that moves rapidly downward from nearly 70 km to less than 25 km altitude during the first few milliseconds after the onset of a lightning discharge. This may be explained as a quasi-electrostatic effect of the extremely rapid removal of previously shielded cloud charge. This moving temperature depression may explain the rapid downward motion displayed by the lower boundary of the sprite-related “elves” phenomenon [Fukunishi et al., 1996].

The second phenomenon is a discontinuity in electron density that propagates downward from roughly 60 km to 30 km along the vertical axis of the simulation during the relaxation period subsequent to the most intense part of the discharge. The discontinuity is seen to propagate for 200 to 300 ms at speeds of approximately 100 km per s. This feature represents a downward propagating wave that extinguishes the process of dissociative attachment in its wake. A one-dimensional, nonlinear attachment wave model suggests that the stationary wave front is maintained under a balance between drift and advection of the charged populations.
6.3 Contributions to Numerical Methods

Adaptations of existing numerical methods are used for all of the calculations required in this effort. The adaptations are for the most part straightforward, with the exception of integrating multiple methods to create the comprehensive simulation described in Section 3.1. The approach adopted is thoroughly original, and quite successful at maintaining a stable marching solution. Moreover, the extension of the flux-corrected transport (FCT) method to a cylindrical geometry, while straightforward to implement, requires a careful examination of the existing method to correctly devise.

A important step in adapting the FCT method for the geometry and boundary conditions of interest is to verify the stability and the conservative property of the method. In so doing, we establish the method as effective and accurate in a whole new class of problems. The same is true for determining the errors in the cylindrical Poisson solution due to discretization and idealized boundary conditions. A careful characterization of such errors must be accomplished for each class of problems addressed. As a result of the examination in Section 3.3.3, the FISHPAK subroutine PWSCYL has been made useful for problems in a cylindrical simulation volume extending from the earth to the ionosphere. Moreover, quantifying the errors inherent in the Poisson solution led to the development of an accurate, cylindrical electric field solution for quasi-neutral environments as described in Section 2.4.3.

6.4 Recommendations for Future Research

Recommendations for future research regarding lightning effects on the lower ionosphere come in two varieties. The first comprise an elaboration of the models developed in this work. The second constitute suggestions for separate efforts inspired by the model results.

The direction of future research into lightning effects on the ionosphere is shaped by the interests of the principal patron of such research, the United States Air Force. As noted in the introduction, unusually intense lightning discharges, of which sprites are a manifestation, have been found to interfere with the unambiguous detection of low-level nuclear blasts. This is an issue of considerable importance in an era when nuclear counter-proliferation is a paramount objective. It is therefore crucial to be able to accurately model the effects of lightning discharges on the overlying ionosphere, and to identify unique aspects of their radiated signatures that might distinguish them from nuclear detonations. The simulation developed here demonstrates that
intense, positive CG lightning generates highly distinctive disturbances in the lower ionosphere. It is likely that unique RF and infrared signatures associated with these disturbances could be observed by satellites and used to resolve the ambiguous cases in nuclear detection. The goal of identifying such signatures provides a context for future research.

6.4.1 Elaboration of Present Models

a. Perform infrared radiance calculations using simulated electron temperatures. The thermal signatures of sprites have not been observed from the earth due to the high temperature of the ambient environment. However, such measurements could be made from space and would be of special interest, as noted above. Rates of electronic excitation of the neutral background are already calculated by the simulation. Additional modeling efforts would be necessary to address radiative emission and absorption in the simulation volume as well as atmospheric absorption along the optical path to a sensor location. Similar calculations have been performed for auroral and dayglow emissions [Wintersteiner et al., 1992], but these calculations have assumed steady state conditions in the emitting region. This is an unacceptable assumption for the study of rapidly evolving lightning-induced effects, given the long radiative lifetimes of some infrared transitions. The development of a dynamic model would be necessary to address infrared emission by sprites.

b. Extend simulation volume to higher altitudes. In Section 4.2 we found that the lightning-induced, transient electric fields observed at altitudes of 90 km and above are likely associated with electromagnetic, and not quasi-electrostatic disturbances [Kelley et al., 1985]. In order to investigate the transionospheric propagation of lightning transients it would therefore be necessary to include the electromagnetic response of the medium. In principle, this involves replacing the Poisson Equation with a combination of Faraday’s and Ampere’s Laws to determine the time varying magnetic and electric fields. In practice, this is a fundamental revision of the simulation, involving the consideration of time scales much shorter than those of interest in this work.

c. Consider sheets of cloud charge. Evidence exists for sprite-associated lightning originating from, and depleting, large horizontal sheets of charge within the extended cloud structure of storms known as mesoscale convective systems [Boccippio et al., 1995]. The horizontal dimensions of a mesoscale convective system are measured in hundreds of kilometers.
Such vast charge sheets would cover the entire horizontal extent of our simulation volume. A straightforward modification of the present cloud charge model would allow us to address this class of charge distributions. In particular, we could allow the cloud charge density to vary only in the vertical dimension, according to a Gaussian distribution, for example.

### 6.4.2 Additional Investigations

a. *Examine horizontal sprite structure.* Sprites exhibit horizontal variations, striations, that are not reflected in the simulation results [Sentman *et al.*, 1995]. One intriguing possibility is that this horizontal structure is associated with pinching instability of the upward electron currents that follow a positive cloud-to-ground discharge. The condition for pinching instability in two counter-streaming electron beams is [Krall and Trivelpiece, 1986, p. 494]

\[
\frac{m_u u_e^2}{\kappa T_e} > \frac{k^2 c^2}{\omega_{pe}^2} + 1
\]  

(6.1)

where \(\omega_{pe}\) is the electron plasma frequency. In terms of the parameters defined in Section 2.7.4, this becomes:

\[
\frac{1}{\eta} > \left( \frac{2\pi c}{L_e} \right)^2 \frac{m_u e_0}{n_0}
\]  

(6.2)

For the conditions associated with 70 km altitude (Section 2.7.5), we find that the left- and right-hand sides have the following magnitudes:

\[
\text{LHS} \sim 0.02; \quad \text{RHS} \sim 10.0
\]  

(6.3)

While it would appear from these numbers that the pinching instability is an unlikely candidate to explain sprite striations, it is nevertheless possible that when we examine the response of the medium on submillisecond time scales, a beam of energetic electrons may escape from middle altitudes before electron-neutral collisions can take effect. The resulting electron population would possess kinetic energy much greater than thermal energy, so that \(\eta \ll 1\), and the pinch condition could be met. However, the simulation is designed for the millisecond time scales of the lightning discharge, and addressing such rapid effects would require a separate formulation.

b. *Perform measurements of RF reflection height above thunderstorms.* Due to the difficulty of making high time resolution, *in-situ* measurements of electron density at ionospheric
altitudes above an active thunderstorm, it is appropriate to seek remote sensing techniques with which to investigate the propagating electron density discontinuity predicted by the simulation in Section 4.4.4. The Rome Laboratory ionospheric channel probe [Myers et al., 1996] can be used to measure propagation characteristics over thunderstorms in conjunction with future Air Force Office of Scientific Research sprite measurement campaigns. The downward propagating density discontinuity should be accompanied by an unmistakable lowering of the high frequency reflection height in the 30 to 60 km altitude region. The frequency range employed by the channel probe is ideally suited to investigating the ten-meter-scale turbulence that appears to be naturally associated with heated volumes in the ionosphere [Kelley, et. al., 1995].
Appendix A

Electron-Neutral Collisional Processes

Table A.1 lists all of the electron-neutral scattering processes incorporated in the calculation of the momentum transfer, excitation, and ionization rates discussed in Section 2.5.1. Collisional cross sections as a function of electron energy were obtained for each process listed [Wadzinski and Jasperse, 1982]. The electron energy loss associated with each of the inelastic processes is shown.

<table>
<thead>
<tr>
<th>Process</th>
<th>Electron Energy Loss, ΔE (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O momentum transfer</td>
<td>6.6E-05 e</td>
</tr>
<tr>
<td>N₂ momentum transfer</td>
<td>3.9E-05 e</td>
</tr>
<tr>
<td>O₂ momentum transfer</td>
<td>3.4E-05 e</td>
</tr>
<tr>
<td>O fine structure, J = 2 to 1</td>
<td>0.0195</td>
</tr>
<tr>
<td>O fine structure, J = 2 to 1</td>
<td>0.0292</td>
</tr>
<tr>
<td>O fine structure, J = 2 to 1</td>
<td>0.0097</td>
</tr>
<tr>
<td>N₂ rotational excitation</td>
<td>0.2</td>
</tr>
<tr>
<td>O₂ rotational excitation</td>
<td>0.2</td>
</tr>
<tr>
<td>N₂ vibrational excitation, v = 0 to 1</td>
<td>0.296</td>
</tr>
<tr>
<td>N₂ vibrational excitation, v = 0 to 2</td>
<td>0.596</td>
</tr>
<tr>
<td>N₂ vibrational excitation, v = 0 to 3</td>
<td>0.899</td>
</tr>
<tr>
<td>N₂ vibrational excitation, v = 0 to 4</td>
<td>1.205</td>
</tr>
<tr>
<td>N₂ vibrational excitation, v = 0 to 5</td>
<td>1.515</td>
</tr>
<tr>
<td>N₂ vibrational excitation, v = 0 to 6</td>
<td>1.829</td>
</tr>
<tr>
<td>N₂ vibrational excitation, v = 0 to 7</td>
<td>2.147</td>
</tr>
<tr>
<td>O electron excitation to ((^2S)) 3s (^3S)</td>
<td>9.52</td>
</tr>
<tr>
<td>O electron excitation to 2p (^4)D</td>
<td>1.96</td>
</tr>
<tr>
<td>O electron excitation to ((^3S)) 3p (^3P)</td>
<td>11.0</td>
</tr>
<tr>
<td>O electron excitation to (2D0) 3s (^3D) (^0)</td>
<td>12.5</td>
</tr>
<tr>
<td>O electron excitation to 2p (^3)P (^0)</td>
<td>15.5</td>
</tr>
<tr>
<td>O electron excitation to ((^3S)) nd (^3D) (^0)</td>
<td>12.75</td>
</tr>
<tr>
<td>O electron excitation to ((^3S)) ns (^3S) (^0)</td>
<td>11.92</td>
</tr>
<tr>
<td>N₂ electron excitation to A (^3\Sigma) (^+)</td>
<td>6.22</td>
</tr>
</tbody>
</table>

Table A.1: Scattering processes and associated energies incorporated in simulation model. Energy loss for momentum transfer collisions are fractions of the incident electron energy. Cont.
<table>
<thead>
<tr>
<th>Process</th>
<th>Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N₂ electron excitation to B²Π_u</td>
<td>7.39</td>
</tr>
<tr>
<td>N₂ electron excitation to W³Δ_u</td>
<td>7.42</td>
</tr>
<tr>
<td>N₂ electron excitation to B¹Σ_u</td>
<td>8.22</td>
</tr>
<tr>
<td>N₂ electron excitation to a¹Σ_u</td>
<td>8.45</td>
</tr>
<tr>
<td>N₂ electron excitation to a¹Σ_u</td>
<td>8.59</td>
</tr>
<tr>
<td>N₂ electron excitation to W¹Δ_u</td>
<td>8.94</td>
</tr>
<tr>
<td>N₂ electron excitation to C³Π_u</td>
<td>11.03</td>
</tr>
<tr>
<td>N₂ electron excitation to E³Σ_u⁺</td>
<td>11.88</td>
</tr>
<tr>
<td>N₂ electron excitation to a''⁴Σ_u⁺</td>
<td>12.25</td>
</tr>
<tr>
<td>N₂ electron excitation to b¹Π_u</td>
<td>12.8</td>
</tr>
<tr>
<td>N₂ electron excitation to b¹Σ_u⁺</td>
<td>13.3</td>
</tr>
<tr>
<td>N₂ electron excitation to N₂⁺ X²Σ_u⁺</td>
<td>13.01</td>
</tr>
<tr>
<td>N₂ electron excitation to N₂⁺ A Π_u</td>
<td>13.19</td>
</tr>
<tr>
<td>N₂ electron excitation to B²Σ_u⁺</td>
<td>15.75</td>
</tr>
<tr>
<td>N₂ electron excitation to D²Π_u</td>
<td>15.71</td>
</tr>
<tr>
<td>N₂ electron excitation to C²Σ_u⁺</td>
<td>20.79</td>
</tr>
<tr>
<td>O₂ electron excitation to a¹Δ_u</td>
<td>0.98</td>
</tr>
<tr>
<td>O₂ electron excitation to b¹Σ_u⁺</td>
<td>1.64</td>
</tr>
<tr>
<td>O₂ electron excitation to A³Σ_u⁺</td>
<td>4.5</td>
</tr>
<tr>
<td>O₂ electron excitation to B³Σ_u⁻</td>
<td>8.4</td>
</tr>
<tr>
<td>O₂ nonionizing dissociation</td>
<td>6.0</td>
</tr>
<tr>
<td>O ionization to O⁺ &quot;S&quot;</td>
<td>13.6</td>
</tr>
<tr>
<td>O ionization to O⁺ &quot;D&quot;</td>
<td>16.9</td>
</tr>
<tr>
<td>O ionization to O⁺ &quot;P&quot;</td>
<td>18.5</td>
</tr>
<tr>
<td>O ionization via (&quot;D&quot;) 3d³P⁺</td>
<td>15.35</td>
</tr>
<tr>
<td>O ionization via (&quot;P&quot;) 3s³P⁺</td>
<td>14.1</td>
</tr>
<tr>
<td>O ionization via (&quot;D&quot;) nd³SPD⁺ Rydberg</td>
<td>16.03</td>
</tr>
<tr>
<td>O ionization via (&quot;P&quot;) nd³PD⁺ Rydberg</td>
<td>17.63</td>
</tr>
<tr>
<td>N₂ ionization to N₂⁺ X²Σ_u⁺</td>
<td>15.58</td>
</tr>
<tr>
<td>N₂ ionization to N₂⁺ A Π_u</td>
<td>16.73</td>
</tr>
<tr>
<td>N₂ ionization to N₂⁺ B²Σ_u⁺</td>
<td>18.75</td>
</tr>
<tr>
<td>N₂ ionization to N₂⁺ D²Π_u</td>
<td>22.0</td>
</tr>
<tr>
<td>N₂ ionization to N₂⁺ C²Σ_u⁺</td>
<td>23.6</td>
</tr>
<tr>
<td>N₂ ionization to N₂⁺ B²Σ_u⁺ Rydberg</td>
<td>17.36</td>
</tr>
<tr>
<td>O₂ ionization to O⁺ &quot;D&quot;</td>
<td>12.1</td>
</tr>
</tbody>
</table>

Table A.1 continued: Scattering processes and associated energies incorporated in simulation model. Energy loss for momentum transfer collisions are fractions of the incident electron energy. Continued.
<table>
<thead>
<tr>
<th>Reaction</th>
<th>Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{O}_2 ) ionization to ( \text{O}_2^+ ) a ( ^4\Pi_u )</td>
<td>16.1</td>
</tr>
<tr>
<td>( \text{O}_2 ) ionization to ( \text{O}_2^+ ) A ( ^2\Pi_u )</td>
<td>16.9</td>
</tr>
<tr>
<td>( \text{O}_2 ) ionization to ( \text{O}_2^+ ) b ( ^4\Sigma_g^- )</td>
<td>18.2</td>
</tr>
<tr>
<td>( \text{O}_2 ) ionization to ( \text{O}_2^+ ) B ( ^2\Sigma_g^+ )</td>
<td>20.0</td>
</tr>
<tr>
<td>( \text{O}_2 ) ionization to ( \text{O}_2^+ ) c ( ^4\Sigma_u )</td>
<td>23.0</td>
</tr>
<tr>
<td>( \text{O}_2 ) ionization to ( \text{O}_2^+ ) a ( ^4\Pi_u )</td>
<td>12.7</td>
</tr>
<tr>
<td>( \text{O}_2 ) ionization to ( \text{O}_2^+ ) A ( ^2\Pi_u )</td>
<td>13.25</td>
</tr>
</tbody>
</table>

Table A.1 continued: Scattering processes and associated energies incorporated in simulation model. Energy loss for momentum transfer collisions are fractions of the incident electron energy.

The processes shown in the table were selected based upon the large relative magnitude of their cross sections. \( \text{O}_2 \) vibrational excitation, for example, was ignored in comparison to \( \text{N}_2 \) vibrational excitation because the associated cross sections were smaller by a factor of 100 over the same energy range. Processes with threshold energies up to the rather large value of 25 eV were included because electron heating to temperatures of several eV was anticipated. Under such circumstances, the high energy tail of the electron velocity distribution is likely to have significant overlap even with high energy cross sections.
Appendix B

Hydration of Small Positive Ions

It has been the approach of researchers studying the chemistry of the D region [Rowe et al., 1974 and Glukhov et al., 1992] to characterize the hydration of NO\(^+\) by the single-body reaction rate

\[ B = 1.0 \times 10^{-43} N^3 \text{ per s} \]  \hspace{1cm} (B.1)

where \( N \) is the total density of neutral particles. This approach does not explicitly account for two of the reactants in the actual three-body process: molecular nitrogen and water vapor. It is reasonable only if two assumptions are satisfied. First, the concentration of molecular nitrogen relative to that of other neutral species must be constant throughout the region, which is true. Second, the water vapor mixing ratio, \( r_w \), must be constant as well. In fact, \( r_w \) varies between roughly 4 ppmm (parts per million, mass) at 50 km and 1 ppmm at 90 km [Grantham et al., 1985]. The use of the single-body rate (2.118) is thus an approximation. Rowe et al. [1974] nevertheless found that the single body approximation yielded successful fits to observations of cluster ion density over the narrow altitude range 70 km to 90 km. Glukhov et al. [1992] apply (B.1) without further justification over the altitude range 30 km to 60 km.

This approach is inadequate for the present simulation because it is our intention to include altitudes essentially down to the ground in the simulation volume. Water vapor mixing ratios as high as 1000 ppmm can be encountered in the lower stratosphere, and we must therefore incorporate the water content of the atmosphere explicitly to obtain a model that is valid over the wide altitude range under consideration. In (2.110) we take into account both the forward and reverse reactions associated with the hydration of NO\(^+\), namely:

\[ NO^+ + H_2O + N_2 \leftrightarrow NO^+ (H_2O) + N_2 \]  \hspace{1cm} (B.2)

with the associated rates:

- hydration (forward reaction) \hspace{1cm} three-body rate \( B_c = 1.6 \times 10^{-40} \text{ m}^6 \text{ per s} \)
- dehydration (reverse reaction) \hspace{1cm} two-body rate \( \lambda_c' = 1.0 \times 10^{-20} \text{ m}^3 \text{ per s} \)

where the dehydration rate, \( \lambda_c' \), given here differs from that appearing in Tables 2.3 and 2.4 because it does not account for the subsequent hydration of \( \text{NO}^+ (H_2O) \). According to Table 2.5, the recombination of cluster ions is relatively insignificant. Thus, if we include (B.2) as the only
hydration reaction pair in our model, we are implicitly assuming that all of the NO\(^+\)(H\(_2\)O) that is created will be available indefinitely for the reverse reaction. In fact, it is transformed into the larger ion NO\(^+\)(H\(_2\)O)\(_2\) even faster than it is created (that is, with an even greater rate coefficient). In order to accurately account for the dehydration reaction in (B.2), we therefore need to correctly model the chain of hydration reactions that lead from NO\(^+\)(H\(_2\)O) to ever larger ions and ultimately to hydronium ions (H\(_3\)O\(^+\)) and HNO\(_2\). We write the pertinent reactions and the associated rates as follows [Rowe et al., 1974].

\[ NO^+ (H_2O) + H_2O + N_2 \rightleftharpoons NO^+ (H_2O)_2 + N_2 \]  
hydration (forward reaction) three-body rate \( \alpha_3 = 1.0E-39 \text{ m}^6 \text{ per s} \)
dehydration (reverse reaction) two-body rate \( \alpha_4 = 1.4E-20 \text{ m}^3 \text{ per s} \)

\[ NO^+ (H_2O)_2 + H_2O + N_2 \rightleftharpoons NO^+ (H_2O)_3 + N_2 \]  
hydration (forward reaction) three-body rate \( \alpha_5 = 2.0E-39 \text{ m}^6 \text{ per s} \)
dehydration (reverse reaction) two-body rate \( \alpha_6 = 1.3E-18 \text{ m}^3 \text{ per s} \)

\[ NO^+ (H_2O)_3 + H_2O \rightarrow H_3O^+ (H_2O)_2 + HNO_2 \]  
hydration (forward reaction) two-body rate \( \alpha_7 = 8.0E-17 \text{ m}^3 \text{ per s} \)

We are able to truncate the hydration chain at (B.5) because there is no significant reverse reaction leading from H\(_3\)O\(^+\)(H\(_2\)O)\(_2\) to NO\(^+\)(H\(_2\)O)\(_3\). We are therefore able to lump H\(_3\)O\(^+\)(H\(_2\)O)\(_2\) together with all subsequent products of hydration and charge exchange. We cease to be concerned with the precise species that compose this class of particles and focus on the processes that return them en masse to the neutral population. The bottom line is that the only means of depleting this entire class of cluster ions is aerosol attachment and recombination. If we use \( n_{c1}, n_{c2}, n_{c3}, \text{ and } n_{c4} \) to denote the densities of NO\(^-\)(H\(_2\)O), NO\(^-\)(H\(_2\)O)\(_2\), NO\(^-\)(H\(_2\)O)\(_3\), and H\(_3\)O\(^+\)(H\(_2\)O)\(_2\) and all subsequent positive ionic products, respectively, we can write the applicable rate equations as:

\[ \frac{d n_{c1}}{d t} = B_C N_{N_2} N_r p - \lambda_C N_{N_2} n_{c1} - \alpha n_m n_{c1} - \alpha_{EC} n_{c1} n_e - \beta N_A n_{c1} \]  

(B.6a)
\[
\frac{d n_{e2}}{dt} = \alpha_3 N_{N_2} N_{e2} n_{e1} - \alpha_4 N_{N_2} n_{e2} - \alpha_5 N_{N_2} N_{e2} n_{e2} + \alpha_6 N_{N_2} n_{e3} - \alpha_n n_{e2} - \alpha_{EC} n_{e2} n_e - \beta N_{A} n_{e2}
\]

\[
\frac{d n_{e3}}{dt} = \alpha_5 N_{N_2} N_{e2} n_{e2} - \alpha_e N_{e2} n_{e3} - \alpha_7 N_{e2} n_{e3} - \alpha_n n_{e3} - \alpha_{EC} n_{e3} n_e - \beta N_{A} n_{e3}
\]  
(B.6c)

\[
\frac{d n_{e4}}{dt} = \alpha_7 N_{e2} n_{e3} - \alpha_n n_{e4} - \alpha_{EC} n_{e4} n_e - \beta N_{A} n_{e4}
\]  
(B.6d)

If we next define the total cluster ion density, \( n_{ct} \), by:

\[
n_{ct} = n_{e1} + n_{e2} + n_{e3} + n_{e4}
\]  
(B.7)

then we can add equations (B.6) together to form a single equation that governs the evolution of \( n_{ct} \):

\[
\frac{d n_{ct}}{dt} = B_{C} N_{N_2} N_{e2} n_p - \lambda_{EC} N_{N_2} n_{e1} - \alpha_n n_{ct} - \alpha_{EC} n_{ct} n_e - \beta N_{A} n_{ct}
\]  
(B.8)

Moreover, we can write the system of rate equations to incorporate (B.8) and to account for all of the processes introduced in Tables 2.3 and 2.4

\[
\frac{d n_e}{dt} = I + G + \lambda n_m N - \alpha_3 (N_{O2})^2 n_e - \alpha_4 N_{O2} n_e - \alpha_5 n_p n_e - \alpha_{EC} n_{e1} n_e - \beta N_{A} n_e
\]  
(B.9a)

\[
\frac{d n_p}{dt} = I + G + \lambda n_m N - B_{C} N_{N_2} N_{e2} n_p - \alpha n_p n_m - \alpha_{EC} n_p n_e - \beta N_{A} n_p
\]  
(B.9b)

\[
\frac{d n_m}{dt} = \alpha_3 (N_{O2})^2 n_e + \alpha_4 N_{O2} n_e - \alpha_5 (n_p + n_{ct}) - \lambda n_m N - \beta N_{A} n_m
\]  
(B.9c)

\[
\frac{d n_{ct}}{dt} = B_{C} N_{N_2} N_{e2} n_p - \lambda'_{EC} N_{N_2} n_{e1} - \alpha n_p n_{ct} - \alpha_{EC} n_{ct} n_e - \beta N_{A} n_{ct}
\]  
(B.9d)

where \( n_e, n_p, n_m, N, N_{O2} \), and \( N_{N_2} \) are electron, small positive ion, small negative ion, aerosol, total neutral, molecular oxygen, and nitrogen densities, respectively. The remaining terms and coefficients are defined in Tables 2.3 and 2.4.

In order to close this system of equations, we must relate \( n_{e1} \) to \( n_{ct} \). To that end, we first note that the summation of equations (B.9) yields:

\[
\frac{d}{dt} \left( n_p + n_{ct} - n_e - n_m \right) = -\beta N_{A} \left( n_p + n_{ct} - n_e - n_m \right)
\]  
(B.10)

so that in the steady state we find:

\[
n_p + n_{ct} - n_e - n_m = 0
\]  
(B.11)

We next seek the steady state ratios of the cluster ion densities. We accomplish this by ignoring the three recombination processes in equations (B.6a), (B.6b), and (B.6c). This simplification is
valid, for we find that the magnitudes of the hydration/dehydration terms in (B.6) vastly exceed the recombination terms. As an example, consider the magnitudes of two of the coefficients in (B.6c), evaluated under the worst-case scenario of 90 km altitude, at which nitrogen density is minimized (CIRA nitrogen density of 5.58E+19 per m³ [COSPAR, 1972]), and electron density is maximized (IRI-90 electron density ~ 1.0E+11 per m³ [Bilitza, 1990]):

$$\alpha_{EC} n_e \approx 0.46 \text{ s}^{-1} \text{ compared to } \alpha_e N_{e2} \approx 72.6 \text{ s}^{-1} \quad \text{(B.12)}$$

where $\alpha_{EC}$ is taken from Table 2.4 (4.6E-12 m³/s) and $\alpha_e$ is taken from (B.4). The rates of all the hydration/dehydration processes are similarly huge compared to the corresponding recombination processes in (B.6). We nevertheless cannot ignore recombination in (B.6d) because it is critical to establishing the equilibrium population of hydration end products with density $n_{e4}$. Making the simplification noted, the steady state forms of (B.6b), (B.6c), and (B.6d) may be solved for the ratios among the cluster ion densities:

$$\beta_2 = \frac{n_{e2}}{n_{e1}} = \frac{\alpha_1 N_{r_w}}{\alpha_3 + \alpha_5 N_{r_w} - \alpha_6 \beta_3} \quad \text{(B.13)}$$

$$\beta_3 = \frac{n_{e3}}{n_{e2}} = \frac{\alpha_5 N_{r_w}}{\alpha_3 + \alpha_6 N_{r_w}} \quad \text{(B.14)}$$

$$\beta_4 = \frac{n_{e4}}{n_{e3}} = \frac{\alpha_7 N_{r_w}}{\beta N_A + \alpha n_m + \alpha_{EC} n_e} \quad \text{(B.15)}$$

If we further assume that $\alpha \approx \alpha_{EC}$, then we can use (B.11) to rewrite (B.15) as:

$$\beta_4 = \frac{n_{e4}}{n_{e3}} = \frac{\alpha_7 N_{r_w}}{\beta N_A + \alpha (n_p + n_{e4})} \quad \text{(B.16)}$$

Since in reality $\alpha_{EC}$ exceeds $\alpha$ by as much as a factor of 10 at high altitudes [Rosen and Hofmann, 1981b], (B.16) overestimates $\beta_4$. Moreover, in order to compute $\beta_4$, we must postulate an altitude profile for $n_{e4}$, since this is not known a priori. On the other hand, $n_p$ is a given quantity in the equilibrium state, as discussed in Section 2.6.2. Figure B.1 illustrates the profile of $n_{e4}$ that is employed in the model. The figure also demonstrates that the quantity of ultimate interest is not strongly sensitive to inaccuracy in either $\alpha$ or $n_{e4}$.

Given the ratios $\beta_2$, $\beta_3$, and $\beta_4$, we rewrite (B.7) as

$$n_{e1} = \delta_H n_{e4} \text{, where } \delta_H = \frac{1}{1 + \beta_2 + \beta_2 \beta_3 + \beta_2 \beta_3 \beta_4} \quad \text{(B.17)}$$
Figure B.1: Influence of recombination coefficient and straw man total cluster ion density on value of $\delta_H$, effective dehydration fraction (B.17). (a) Comparison of actual $\delta_H$ employed in model with two alternatives. Bottom curve shows profile corresponding to alternate cluster ion density shown in Part b. Broken curve shows profile for choice of electron-cluster ion recombination coefficient, $\alpha_{EC}$, instead of ion-ion recombination coefficient, $\alpha$, in (B.15). Comparison demonstrates that $\delta_H$ is insensitive to the choice of cluster ion profile and recombination coefficient. (b) Solid curve shows straw man total cluster ion density employed in chemical model (B.16). Broken curve is alternate density, used to show that precise choice of profile is not critical to the eventual calculation of $\delta_H$. 
which is the result we are seeking. Using (B.17), we recast (B.9) as:

\[ \frac{dn_e}{dt} = I + G + \lambda n_m N - \alpha_3 (N_{O2})^2 n_e - \alpha_D N_{O2} n_e - \alpha_E n_p n_e - \alpha_{EC} n_{et} n_e - \beta N_A n_e \]  

(B.18a)

\[ \frac{dn_p}{dt} = I + G + \delta_H \lambda'_{C'} N_{N_2} n_{et} - B_C N_{N_2} N_{W} n_p - \alpha n_p n_m - \alpha_E n_p n_e - \beta N_A n_p \]  

(B.18b)

\[ \frac{dn_m}{dt} = \alpha_3 (N_{O2})^2 n_e + \alpha_D N_{O2} n_e - \alpha n_m (n_p + n_{et}) - \lambda n_m N - \beta N_A n_m \]  

(B.18c)

\[ \frac{dn_{et}}{dt} = B_C N_{N_2} N_{W} n_p - \delta_H \lambda'_{C'} N_{N_2} n_{et} - \alpha n_m n_{et} - \alpha_{EC} n_{et} n_e - \beta N_A n_{et} \]  

(B.18d)

which are precisely the equations used in the simulation, (2.110). The only differences are cosmetic, with

\[ n_{et} \leftrightarrow n_e \quad \text{and} \quad \delta_H \lambda'_{C'} \leftrightarrow \lambda_C \]  

(B.19)

As suggested in Section 2.6.1, we see that the density \( n_e \), by virtue of its equivalence with total cluster ion density \( n_{et} \), does encompass all possible cluster ions. The ultimate result of our careful treatment of positive ion hydration is that the dehydration reaction rate in (B.18b) and (B.18d) has been reduced by the factor \( \delta_H \). This modification accounts for the fact that NO\(^+\)(H\(_2\)O), the direct product of hydrating the small positive ion NO\(^+\), may quickly evolve into even larger ions and thereby become unavailable for the reverse reaction. We see from Figure B.1 that this effect can be quite important at low altitudes.

Finally, we also point out that due to their fast rates, the hydration process will quickly reach equilibrium on the millisecond time scales of interest in the simulation. As an example, the representative reaction pair (B.4), at the representative altitude of 50 km (CIRA nitrogen density of 1.74E+22 per m\(^3\) [COSPAR, 1972]), exhibits a relaxation time constant of

\[ \frac{1}{\alpha_e N_{N_2}} = 44.3 \mu s \]  

(B.20)

which is very fast on the time scales of interest in this problem. At the low altitudes for which the hydration correction is significant, neutral density is even greater and (B.20) yields an even shorter time. The fact that equilibrium is achieved so quickly permits us to use the equilibrium value for \( \delta_H \) in the dynamic rate equations (B.18).
Appendix C

Monopolar Versus Dipolar Cloud Charge Models

In Section 2.3 we discussed two models for the distribution of charge within a thundercloud. The dipolar model is generally accepted by the meteorological community as a reasonable representation of reality. It consists of a positive charge center located vertically above a negative charge center. The alternative monopolar model, put forth here and in Pasko et al. [1995], consists of only the positive charge center. The monopolar model has the virtue of simplicity, and for our principal purpose of studying the effects of positive cloud-to-ground lightning discharges, it accurately captures the physics in the region of interest. It is the goal of this appendix to demonstrate that no loss of generality results from the use of the monopolar model.

To this end, we compare the simulation results obtained for identical lightning discharges under the two different cloud charge models. The discharge considered lowers a total charge of +100 C to ground with an e-folding time of 1.25 ms. The cloud electrification occurs over a 5 s period. The positive charge center is in both cases located at an altitude of 15 km on the simulation axis. The negative charge center is at an altitude of 8 km, consistent with Krehbiel [1986, p. 96]. As illustrated in Figure 2.5, the primary difference in the stimulus to the system between the two models is that the dipolar model entails an equal and opposite growth of cloud charge in the two charge centers during the electrification period, so that the net charge in the low altitude region of the simulation volume remains at zero. This contrasts with the monopole model, in which the total low altitude charge grows directly with the cloud electrification.

Because the stimulus to the system is non-neutral under the monopolar cloud model, we might anticipate larger electric field values and greater electron heating when using this model. While this is true to a limited degree during the electrification period, it is certainly not true after the onset of the discharge. This fact is demonstrated by Figure C.1, which shows corresponding electron temperatures throughout the simulation volume 3 ms after the onset of the discharge. There is no discernible difference between these temperature fields anywhere in the simulation.
Figure C.1: Comparison of electron temperatures obtained for identical lightning discharges under the monopolar and the dipolar cloud charge models. The lightning discharge lowers +100 C from the positive charge center at 15 km altitude, with an e-folding time of 1.25 ms. Electrification lasts 5 s. Plots are obtained at 3 ms after onset of discharge.
volume except directly beneath the thundercloud itself. Even the difference in that small region is not dramatic. In any case, the region of the simulation volume below cloud height is not of interest in the present study of lightning effects on the lower ionosphere.

The small magnitude of the difference in simulation results after the onset of the discharge is emphasized by the electron temperature cut along the simulation axis shown in Figure C.2a. Again, the only discrepancy noted is in the immediate vicinity of the lower cloud charge center. On the other hand, Figure C.2b does show that electron heating during the electrification period is greater by up to 20% under the monopolar cloud model than under the dipolar model. It is also clear that the difference vanishes with the onset of the discharge.

The reason for the lack of any significant difference between results obtained using the two different cloud charge models is that the medium is able to mostly neutralize low altitude charge buildup in any configuration during the lengthy electrification period. The neutralization is not perfect, as the exponential time dependence of the cloud electrification is especially pernicious, but it is good enough to prevent significant heating during the electrification period at all altitudes above the thundercloud. The discharge, on the other hand, is so rapid that it outpaces the medium's response at all altitudes below 70 km. As a result, large electric fields and significant electron heating result during the discharge period.

The bottom line is that the system is able to maintain an essentially field-free quasi-equilibrium throughout the electrification. The event that has the strongest effect on the fields and temperatures is the discharge itself, which originates from the positive charge center—regardless of the cloud charge model used. Thus, the major stimulus to the system is identical under the two models, and we obtain nearly identical results. We conclude that the results we obtain by using the monopolar cloud model apply with great accuracy to the dipolar case as well.
Figure C.2: Comparison of simulation results obtained using the monopolar and dipolar cloud charge models for discharge parameters of Figure C.1. Broken curves refer to monopolar model, solid curves to dipolar model. (a) Altitude profile of electron temperature along simulation axis at 3 ms after onset of discharge. (b) Time series of electron temperature at altitude 56.4 km on simulation axis.
Bibliography


Scott, W. B., Composite satellite to detect covert nuclear tests, *Aviation Week & Space Technology*, 48, 27 Nov 95.


Wilson, C. T. R., The acceleration of \( \beta \)-particles in strong electric fields such as those of thunderclouds, Cambridge Phil. Soc. Proc., 22, 534, 1925.


THESIS PROCESSING SLIP

FIXED FIELD: iill. __________________ name __________________

index ____________ biblio ____________

► COPIES: Archives Aero Dewey Eng Hum

Lindgren Music Rotch Science

TITLE VARIES: □ __________________________

NAME VARIES: □ __________________________

IMPRINT: (COPYRIGHT) __________________

► COLLATION: 221p

► ADD. DEGREE: ____________ ► DEPT.: ____________

SUPERVISORS: __________________________

________________________

________________________

________________________

________________________

________________________

________________________

________________________

________________________

NOTES:

cat'r: __________________ date: __________

► DEPT: E.E. page: J167


► NAME: ERNSTMEYER, James