Structure of the Polar Electrojet Antenna

by

Ryan James Riddolls

S.B., Electrical Engineering and Computer Science, MIT (1997)

M.Eng., Electrical Engineering and Computer Science, MIT (1999)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

June 2003

Author . . . .
Department of Electrical Engineering and Computer Science
May 22, 2003

Certified by . . . .
Min-Chang Lee
Leader of Ionospheric Plasma Research Group
Thesis Supervisor

Certified by . . . .
Ronald R. Parker
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Structure of the Polar Electrojet Antenna

by

Ryan James Riddolls

Submitted to the Department of Electrical Engineering and Computer Science
on May 22, 2003, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Field experiments were performed in Gakona, Alaska in August and November 2002. The ionospheric conductivity was periodically perturbed using amplitude-modulated radiation from the HAARP HF transmitter (1 MW power, 14 dB gain, 3.3–5.8 MHz carrier, 0.1–40 kHz modulation). The conductivity perturbations lead to perturbations to the natural flow of electrojet current in the lower ionosphere, resulting in ELF/VLF radiation at the modulation frequency. Measurements of the radiation, along with analytic and numerical models, suggest that a vertical loop with a scale size of 10 km is the dominant current structure excited during the experiments.

Thesis Supervisor: Min-Chang Lee
Title: Leader of Ionospheric Plasma Research Group

Thesis Supervisor: Ronald R. Parker
Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

Professor Min-Chang Lee has been my research supervisor throughout my ten years as an MIT student. He has supported and nurtured my development for an extraordinary period of time. I thank him with all my heart for his commitment and generosity at every step of the journey.

I would also like to thank Professors Ronald Parker and Abraham Bers who have overseen my progress through the doctoral program over the last four and a half years. I am grateful to Professor Parker for taking numerous risks during that time, such as assigning me to unconventional teaching duties, and helping supervise a thesis with uncertain prospects for success. Professor Bers has gone beyond the call of duty as my qualifying exam chair, and later thesis reader. I thank him for his timely and comprehensive feedback on my research paper, thesis proposal, and draft of this document.

Lastly, I would like to thank Daniel Hernandez, who along with Professor Lee and I, spent many long hours in cold conditions monitoring the health of the data acquisition process.

Contents

List of Figures

11

List of Tables

Chapter 1

Introduction

1.1   Thesis overview

This thesis investigates the structure of ionospheric currents excited during heating experiments in the high-latitude ionosphere. Heating of the ionosphere is accomplished using a ground-based radio transmitter in the low HF (3–10 MHz) range of frequencies. A significant amount of heating occurs in the C and D regions of the ionosphere (50–90 km altitude), resulting in the modification of the electron conductivity within a volume with a linear scale of 10–20 km. The resulting discontinuity in the conductivity perturbs the natural flow of horizontal electron current in the polar region. If the transmitted HF power is periodically modulated, then the pattern of perturbed current radiates at the modulation frequency. Due to the large scale size of the perturbed current, this scheme offers an effective way to generate electromagnetic radiation at frequencies in the ELF/VLF range (below 30 kHz). This scheme is shown pictorially in Figure 1-1.

ELF/VLF radiation is essential for communication with receiving stations located deep underground or underwater. For example, shown in Figure 1-2 is the attenuation of an electromagnetic wave in sea water versus the depth of the receiving station and frequency. From the figure, it is clear that one needs to operate below 10 kHz to remain in contact with the receiving station. However, massive ground-based ELF/VLF transmitters suffer from economic and geopolitical barriers. An ionosphere-
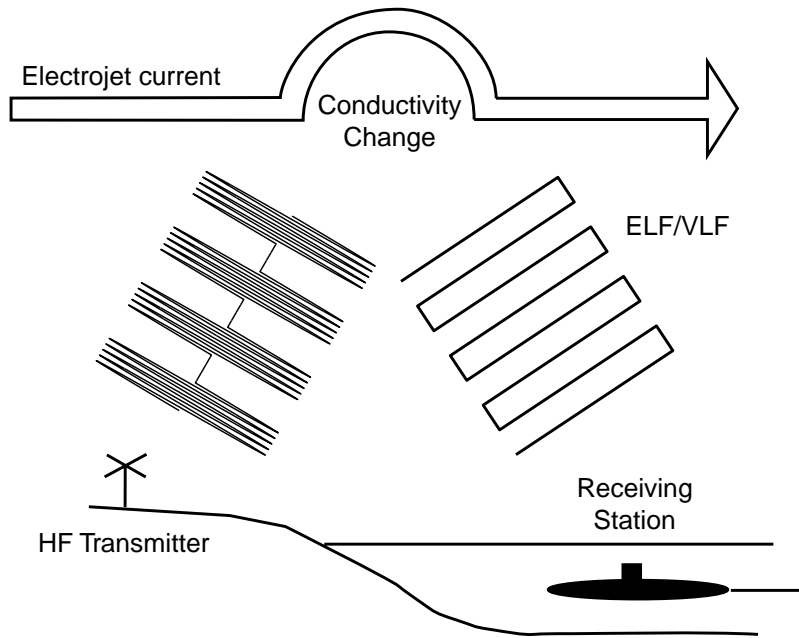
Figure 1-1: ELF/VLF wave generation scheme.
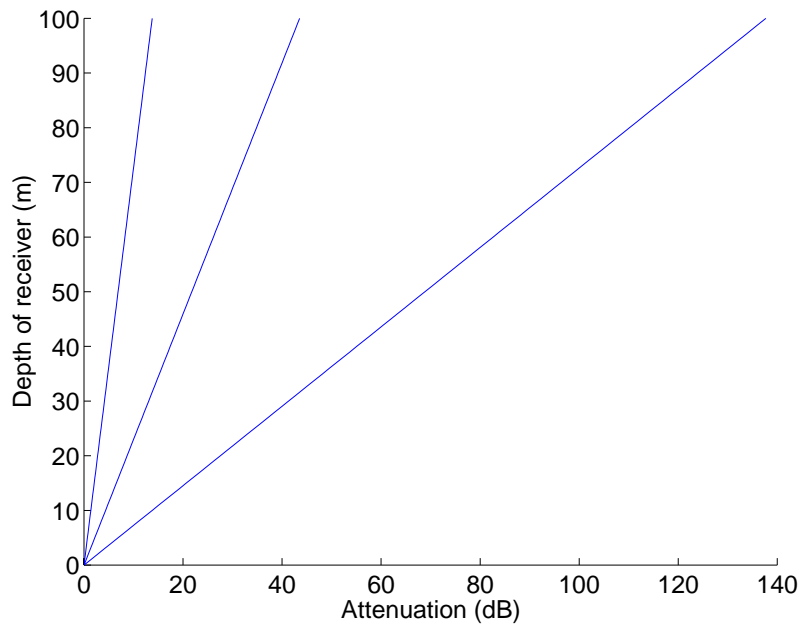
Figure 1-2: Attenuation of an electromagnetic wave versus depth of the receiving station in salt water.

based system is much more flexible in that it can be excited with a comparatively small ground-based HF transmitter, and it does not need to be retuned for operation at different ELF/VLF frequencies.

However, the tradeoff is that the operator has limited knowledge of the structure of the antenna currents generated during HF heating. Specifically, the pattern of perturbed current cannot be determined directly due to a lack of *in situ* instruments. Thus one must be content with observing the radiation on the ground and extrapolating backwards to the source region. In this thesis, the author attempts to deduce the structure of the currents based on measurements of polarization, magnitude, and phase of ELF/VLF radiation during an ionospheric heating experiment using the HAARP facility in eastern Alaska. Two simple theories of antenna current are considered [*Stubbe and Kopka*, 1977, and *Stubbe et al.*, 1982] and are found to be insufficient to explain the observed radiation. The theories are extended by the author to explain the experimental data.

The introductory chapter briefly discusses the ionospheric environment and outlines ionospheric heating. Chapter 2 reviews the kinetics of plasma heating. The Boltzmann equation is analyzed assuming elastic binary collisions and stationary molecules. This approach is carried until it becomes obviously necessary to include molecular motion and inelastic collisions. Effects such as molecular recoil and rotational excitation of molecules are considered. The formulation is continued until we arrive at expressions for the distribution function in terms of the power of the injected HF radiation.

Chapter 3 solves for the distribution function of the heated plasma and then determines the bulk parameters of temperature, conductivity, and current. Two distinct limits for the distribution function are considered. The first limit assumes the modulation is slow compared to the thermal response time of the ionosphere, and thus the ionosphere is in thermal steady state. The other limit is the opposite case, where the modulation frequency is fast compared to the response time and the electron distribution function exhibits small perturbations around an average value. The current is described in several levels of the theoretical detail. The first level

describes the theory of *Stubbe et al.* [1982], where the polarization of the modified plasma volume is ignored. The second level is the theory of *Stubbe and Kopka* [1977], which gives a two-dimensional analytic model for static polarization. The third level is the author's extension of the second level to three dimensions. The fourth level describes the author's numerical solution to Maxwell's equations.

Chapter 4 presents data from heating experiments in Alaska. Measurements of the radiation polarization ellipse and magnitude/phase are used to test predictions of the theory. It is found that the author's third- and fourth-level theories are necessary to explain the observed characteristics of the radiation.

Chapter 5 summarizes the thesis and assesses the results of the experiments. Some suggestions for future work are provided.

1.2   The plasma environment

The solar wind consists of hot, fully-ionized hydrogen and helium ($n \approx 10^7$ m^{-3}, $T_e \approx$ 10 eV) expanding outward from the sun at approximately 400 km/s. The magnetic diffusion time $\tau_m = \sigma \mu_0 L^2$ is thousands of years, which is long compared to the 4 days it takes for the solar wind to travel from the sun to the earth. Thus the sun's magnetic field is "frozen in" to the expanding plasma. A top view of the sun's magnetic field is shown in Figure 1-3. The magnetic field spirals outward with the velocity of the solar wind, and rotates with the 27-day period of the sun, resulting in the spiral pattern shown in the figure. There is also vertical structure. As the solar wind expands, the magnetic field is stretched into a thin disc. Field lines above and below the ecliptic plane are oppositely directed, and therefore separated by a current sheet. This current sheet, as shown in Figure 1-4, undulates with the 27-day rotational period of the sun, since the magnetic and geographic poles of the sun do not coincide. The inset in Figure 1-4 shows the field structure in a vertical plane. As the current sheet undulates, the sun's magnetic field alternates upward and downward with respect to the ecliptic at the location of the earth.

Let us consider the case of downward-directed field lines at the location of the
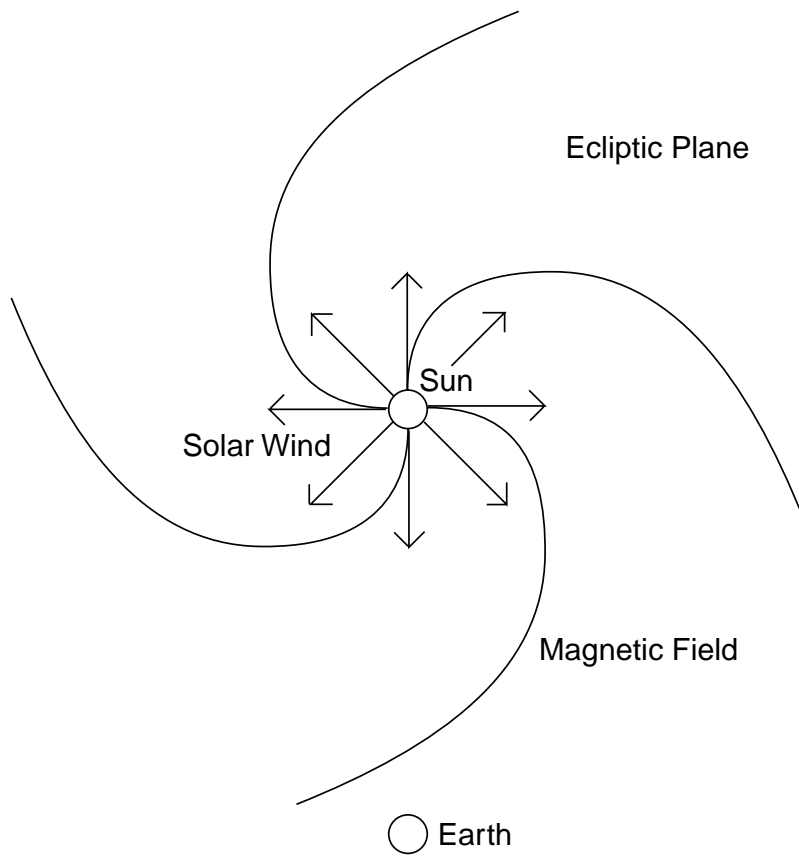
Figure 1-3: Top view of the sun's magnetic field.
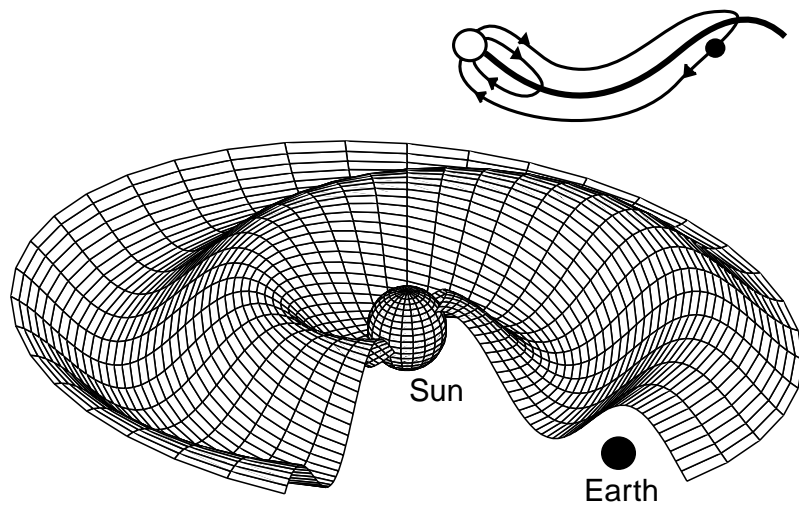
Figure 1-4: Current sheet.
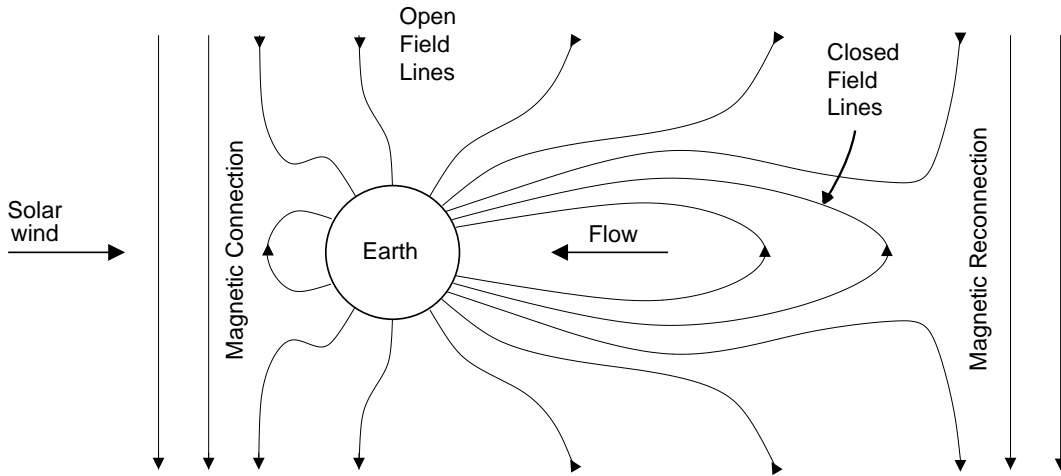
Figure 1-5: A simple model of the earth's magnetosphere.

earth. Shown in Figure 1-5 is the earth's magnetosphere in such a case, where downward-directed magnetic field lines from the sun approach the earth. The sun's field connects with the earth's field on the sunward side of the earth, and the whole structure is dragged away from the sun until reconnection occurs in the earth's magnetotail. The magnetotail is a distorted dipole field, and so $\nabla \times \mathbf{B}$ is nonzero there. A current must exist coming out of the plane of the figure, and thus a $\mathbf{J} \times \mathbf{B}$ force will drive plasma flow back towards the earth. Of particular interest here is the effect of plasma flow across the earth's magnetic field lines in the magnetosphere, as shown in Figure 1-6. The closed dipolar field lines around the side of the earth experience a transverse flow in the sunward direction, resulting in an outward-directed (evening sector) or inward-directed (morning sector) electric field, according to the ideal MHD "Ohm's law" $\mathbf{E} \approx -\mathbf{v} \times \mathbf{B}$. However, the conductivity along the field lines is high, so field-aligned currents $\mathbf{J}_\parallel$ transport this electric field downward to the ionosphere where it is directed in a general northern direction (evening sector), or southern direction (morning sector). These electric fields drive plasma flow in the polar F region of the ionosphere (above 150 km), again according to Ohm's law. Since the electric fields change sign between the evening and morning sectors, the resulting plasma convection pattern has a two-cell structure. An example of a two-cell structure deduced from
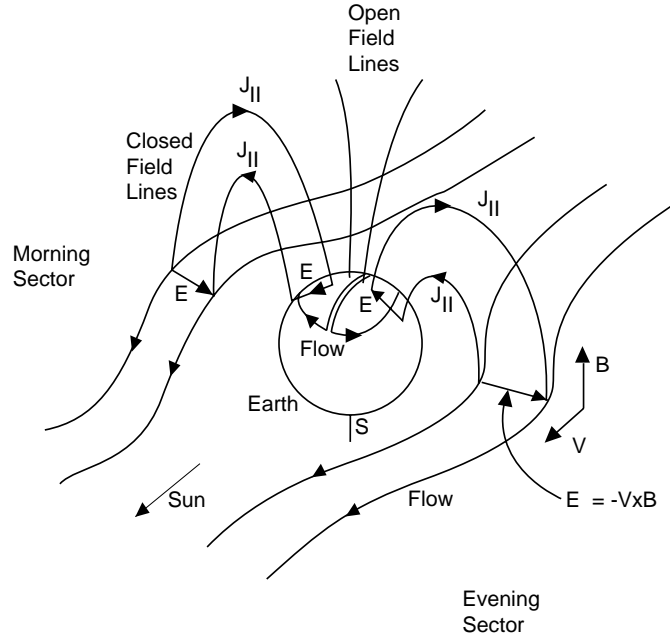
Figure 1-6: Dynamo action.

satellite magnetic measurements is shown in Figure 1-7 (from the CEDAR/TIMED database). The electric fields also map down to the lower regions of the ionosphere (below 150 km). In the lower layers, collisions are more frequent, leading to independent behaviour of the electrons and ions, and hence allowing electric current. In the E region (90-150 km), the crossfield (Pedersen) electron conductivity is quite high in a rather small altitude range (see Figure 1-8), leading to a narrow layer of current referred to as the polar electrojet. It should be noted that this electrojet mechanism is quite distinct from the equatorial electrojet, which relies on neutral F region wind as a drive rather than magnetospheric solar wind dynamo action.

While the polar electrojet is centered in the E region, small amounts of current also flow in the C and D regions (50-90 km). It is at these altitudes where the conductivity perturbation due to ionospheric heating can significantly disturb the natural electric current. However, our ability to understand what happens in the C and D regions during the heating experiments is limited by our knowledge of the electron density profile. The C and D region electron densities cannot be investigated using standard reflectometry or Thomson scattering methods, due to low density. Rocket

23
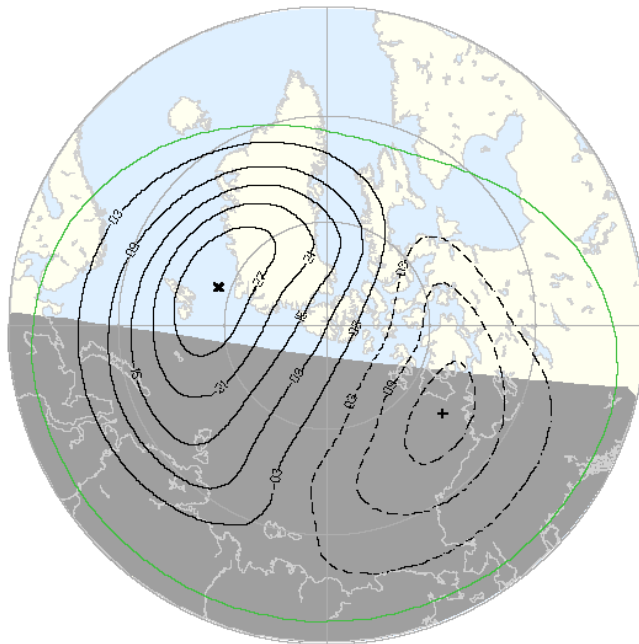
Figure 1-7: Two-cell pattern of auroral F region plasma convection.
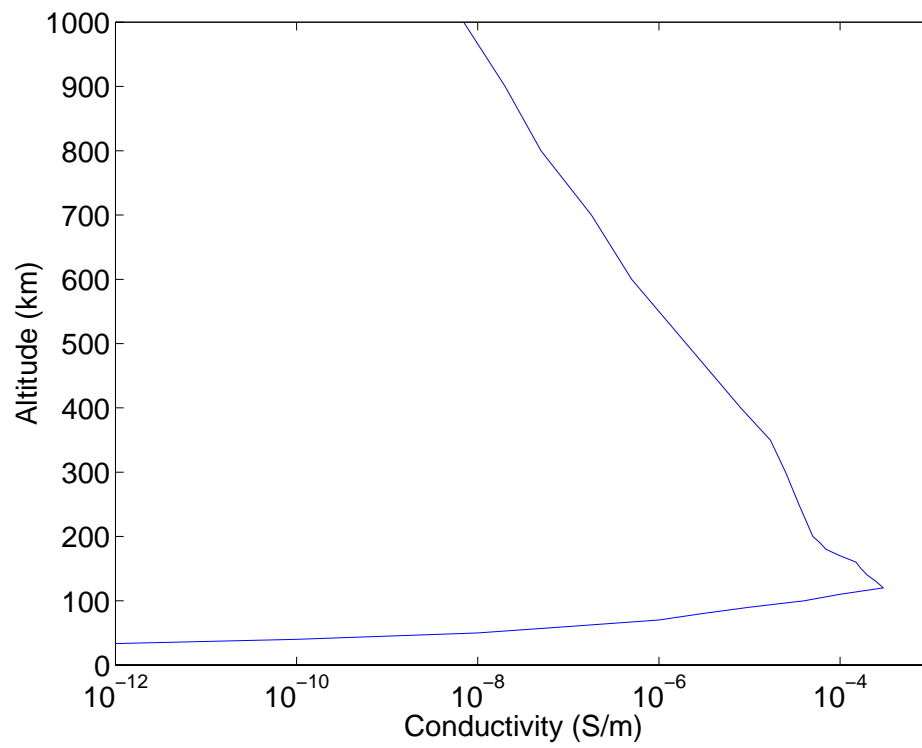
Figure 1-8: Crossfield (Pedersen) electron conductivity.
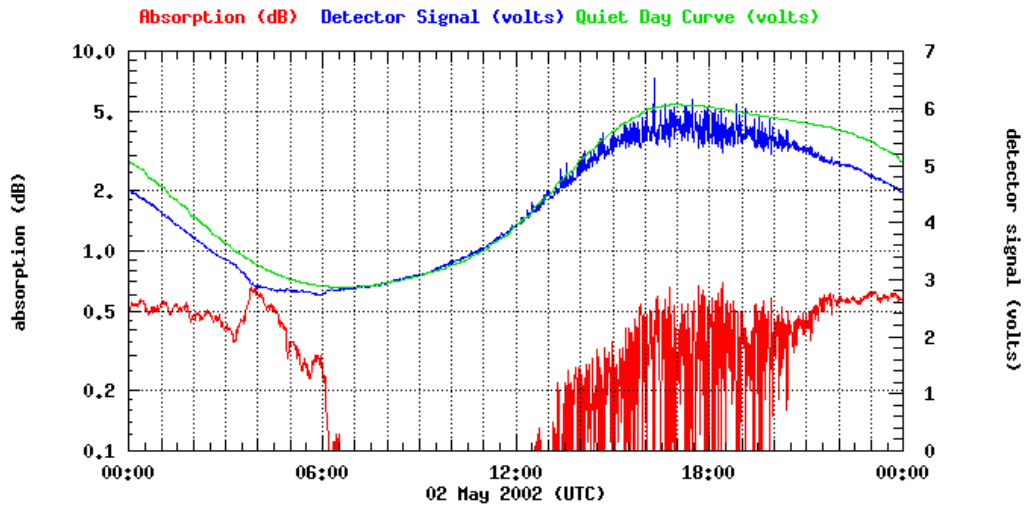
Figure 1-9: Riometer data.

or satellite measurements are possible, but are not continuous. Thus one usually settles for a height-integrated measurement technique, such as riometry. A riometer is a ground-based zenith-pointing receiver which monitors the level of galactic noise near a frequency of 30 MHz. The noise originates from various extra-terrestrial sources, and varies from one hour to another as the Earth rotates. If there was no ionospheric absorption, then the level of noise over a period of one day would be a repeatable function of the time of day. After many observations, one can develop a "quiet day" curve which represents the maximum level of noise that is received at a given hour of the day. By comparing the current noise level to the quiet day curve, the current level of ionospheric absorption can be determined. Generally speaking, if the signal level is large, then the height-integrated D region plasma density is low, whereas if the signal is small, then the height-integrated density is large. Thus the riometer gives some idea of the lower ionospheric density. An example is shown in Figure 1-9 (from the HAARP instrument suite). The top curve is the quiet-day curve, the middle curve is the measured 30 MHz signal level and the bottom curve is the deduced absorption.

There have been attempts to associate electron density profiles with different levels of galactic noise absorption. Rocket-based instruments have been flown through the

25
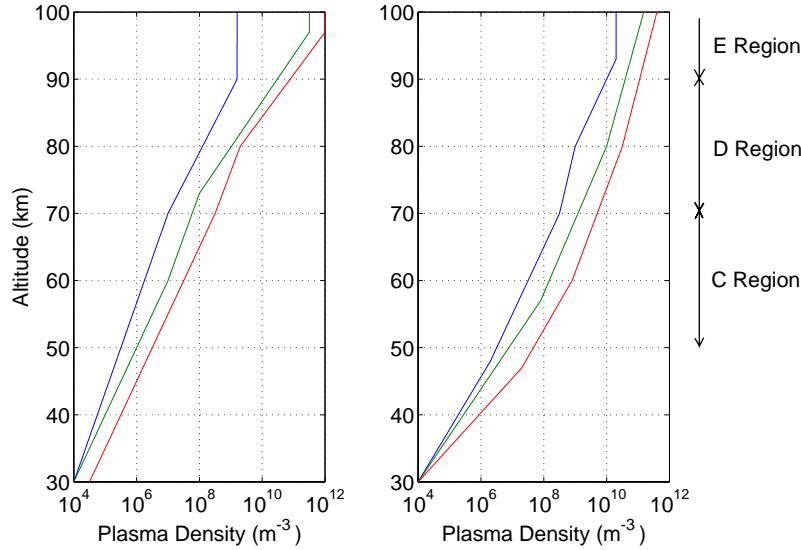
Figure 1-10: Left panel, nighttime electron density profiles for (left to right) 0.1, 0.6, and 1.8 dB of absorption. Right panel, daytime profiles for (left to right) 0.1, 0.8, and 3.2 dB of absorption.

C/D region of the high-latitude ionosphere at different times of day and during different galactic noise absorption conditions [*Jespersen*, 1966] in an attempt to produce a "catalog" of profiles indexed by time of day and absorption level. A simple catalog of this sort was presented by *Barr and Stubbe* [1984], displayed in Figure 1-10. Shown are six electron density profiles representing typical nighttime and daytime profiles in conditions of low, medium, and high absorption of 30 MHz galactic noise. These profiles will be adopted when we discuss modification of the lower ionosphere. In this study we will not use the often-cited mid-latitude profiles given by *Gurevich* [1978].

For completeness, we mention the neutral profile, as it plays an important role in the loss processes. In the 30–100 km range of altitudes, the neutral density gently rolls off with increasing altitude, maintaining the familiar 4:1 nitrogen-to-oxygen ratio. The profile is plotted in Figure 1-11 [*Johnson*, 1961].

26
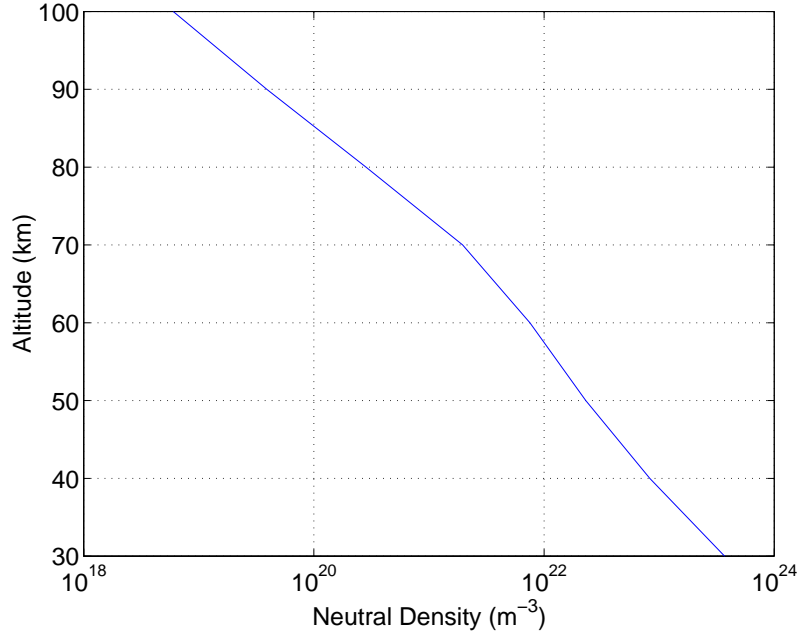
Figure 1-11: Neutral density profile.

1.3 HF modification

Powerful HF waves launched from ground-based transmitting stations (referred to as "heaters") can significantly modify the lower ionosphere. The electron temperature increases, affecting elastic and inelastic electron-neutral collision frequencies, which results in a change to the plasma conductivity. Local changes to the plasma conductivity in the presence of solar wind dynamo-produced electric fields will lead to perturbed electric currents. In particular, amplitude-modulated HF transmissions will lead to "antenna" radiation at the modulation frequency by the perturbed currents.

The mechanism was first demonstrated at Gor'kii in the USSR [*Getmantsev et al.*, 1974] using two carrier frequencies, resulting in beat-wave modulation of the conductivity. The effect was confirmed in the polar region at Monchegorsk [*Kapustin et al.*, 1977]. Later experiments in the 1980s with the EISCAT transmitter at Tromsø, Norway [*Stubbe et al.*, 1981 and 1982, *Rietveld et al.*, 1986 and 1987], are the most comprehensive to date, having produced much frequency and time domain analysis of the radiating characteristics. However, it soon became clear that the major limi-

tation to achieving large temperature (and hence conductivity) changes in the lower ionosphere was due to a process known as self-absorption.

As the ionosphere heats up during the heater pulse, the effective electron-neutral collision rate increases, leading to increased attenuation of the injected HF wave. Thus the process of ionospheric heating is self-reinforcing such that lower ionospheric regions quickly run away to high temperatures while higher altitudes do not get significantly heated. This problem of self-absorption has the obvious consequence that only low altitudes can be heated to high temperatures [*Kotik and Itkina*, 1998]. Unfortunately, conductivity modulation is a much more effective process at higher altitudes, as it scales with plasma density, and thus the full potential of ionospheric heating cannot be realized.

Some failed approaches at a solution include heating the ionosphere for short periods of time, such that the lower layers do not have sufficient time to heat up, and also launching waves from the topside of the ionosphere to access higher altitudes directly. More practical recent approaches have looked at rapidly scanning the transmitter beam angle to heat a larger horizontal extent of plasma [*Papadopoulos et al.*, 1990], or to excite Cerenkov radiation [*Papadopoulos et al.*, 1994]. In addition, scanning allows one to use the ELF/VLF source as a probe for underground tomography applications, and thus demands good control of the polarization of the radiation [*Milikh et al.*, 1999]. Realizing these endeavors requires a fundamental understanding of the nature of the current structure which is excited during an ionospheric heating experiment, and thus the motivation for pursuing the current work.

Chapter 2

The Kinetics of Ionospheric Modification

The theory presented in this chapter is the author's interpretation of the kinetics described by *Allis* [1956], *Shkarofsky et al.* [1966], and *Gurevich* [1978]. The aim is to provide a concise, but meaningful motivation for Equation (3.1), which is the point of departure for the author's own calculations in Chapters 3 and 4.

In the ionospheric C and D regions (50–90 km), the frequency of an HF wave (3–10 MHz) is much larger than the typical electron plasma frequency (100 kHz). Thus we can safely ignore the dynamics of ions. Furthermore, inhomogeneity can be ignored because the mean free path (1–10 cm, for plasma at $T_e = 180$ K and $N = 10^{21}$ m^{-3}) is short compared to the smallest inhomogeneity scale length, the HF wavelength (100 m). In other words, electrons will make many collisions before moving a significant fraction of an HF wavelength. Thus we investigate the homogeneous Boltzmann equation for the electrons:

$$\frac{\partial f}{\partial t} - \frac{e}{m}(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f}{\partial \mathbf{v}} + S = 0. \tag{2.1}$$

Here, f is the electron distribution function, and S is the Boltzmann collision integral, which describes the change in f due to collisions. The collisions are dominantly electron-neutral. The interaction forces are short range, and therefore we consider

only binary collisions.

The distribution f is a function of three velocity variables and one time variable. In spherical coordinates one can write $f(\mathbf{v}, t) = f(v, \theta, \phi, t)$. A procedure for handling such dependencies is to expand f in a series of orthogonal basis functions of one or more of the independent variables. In our case, we will expand two dimensions of velocity space (the polar and azimuthal angles) in spherical harmonics, and the one dimension of time in Fourier series.

2.1 Spherical harmonic expansion

2.1.1 Expansion to first order

In view of the spherical geometry of the collision integral, we consider an expansion of f in spherical harmonics:

$$
\begin{aligned}
f(v, \theta, \phi) &= \sum_{lms} f_{lms}(v) Y_{lms}(\theta, \phi) & (2.2) \\
&= \sum_{l=0}^{\infty} \sum_{m=0}^{l} \sum_{s=0}^{1} f_{lms}(v) Y_{lms}(\theta, \phi). & (2.3)
\end{aligned}
$$

For notational simplicity, we do not denote the time dependence of f. The spherical harmonic Y_{lms} is defined by

$$
Y_{lms}(\theta, \phi) = P_l^m(\cos \theta)(\delta_{0s} \cos m\phi + \delta_{1s} \sin m\phi). \tag{2.4}
$$

Here, θ is the angle between $\mathbf{v}$ and $\mathbf{B}$, ϕ is the azimuthal coordinate, and the Legendre polynomials $P_l^m(\cos \theta)$ are defined by

$$
P_l^m(\cos \theta) = \frac{(2l)! \sin^m \theta}{2^l l! (l - m)!} \left[\cos^{l-m} \theta - \frac{(l - m)(l - m - 1)}{2(2l - 1)} \cos^{l-m-2} \theta + ... \right]. \tag{2.5}
$$

To start, we will expand f to order $l = 1$. This confines the current treatment to cases where f is only weakly anisotropic. The requirements for weak anisotropy are not easily established at the beginning, but the validity of the approximation can be

verified *a posteriori* (see Section 3.1.1). We proceed by expanding:

$$f(v, \theta, \phi) \approx \sum_{l=0}^{1} \sum_{m=0}^{l} \sum_{s=0}^{1} f_{lms}(v) Y_{lms}(\theta, \phi) \tag{2.6}$$

$$= f_{000} Y_{000} + f_{110} Y_{110} + f_{111} Y_{111} + f_{100} Y_{100} \tag{2.7}$$

$$= f_{000} + f_{110} \sin\theta \cos\phi + f_{111} \sin\theta \sin\phi + f_{100} \cos\theta. \tag{2.8}$$

Here we have written out the $l = 0$ and $l = 1$ spherical harmonics explicitly. These harmonics have a useful physical interpretation. The Y_{000} harmonic is isotropic, whereas the other three are the directional cosines of a vector in three-dimensional velocity space. By adopting the following definitions:

$$f_0 = f_{000} \tag{2.9}$$

$$\mathbf{f}_1 = f_{110}\hat{\mathbf{x}} + f_{111}\hat{\mathbf{y}} + f_{100}\hat{\mathbf{z}}, \tag{2.10}$$

we can interpret Equation (2.8) as the sum of isotropic and directional components:

$$f(\mathbf{v}) \approx f_0(v) + \mathbf{f}_1(v) \cdot (\mathbf{v}/v). \tag{2.11}$$

We now insert the first order expansion (2.11) into the Boltzmann equation (2.1). The procedure is written out here term by term. The time derivative term is

$$\frac{\partial}{\partial t}\left(f_0 + \mathbf{f}_1 \cdot \frac{\mathbf{v}}{v}\right) = \frac{\partial f_0}{\partial t} + \frac{\mathbf{v}}{v} \cdot \frac{\partial \mathbf{f}_1}{\partial t}. \tag{2.12}$$

The electric field term is

$$\mathbf{E} \cdot \frac{\partial}{\partial \mathbf{v}}\left(f_0 + \mathbf{f}_1 \cdot \frac{\mathbf{v}}{v}\right) = \mathbf{E} \cdot \left[\frac{\mathbf{v}}{v}\frac{\partial f_0}{\partial v} + \frac{\partial}{\partial \mathbf{v}}\left(\mathbf{f}_1 \cdot \frac{\mathbf{v}}{v}\right)\right] \tag{2.13}$$

$$= \mathbf{E} \cdot \left[\frac{\mathbf{v}}{v}\frac{\partial f_0}{\partial v} + \frac{\mathbf{f}_1}{v} \cdot \frac{\partial \mathbf{v}}{\partial \mathbf{v}} + \frac{\partial}{\partial \mathbf{v}}\left(\frac{\mathbf{f}_1}{v}\right) \cdot \mathbf{v}\right] \tag{2.14}$$

$$= \mathbf{E} \cdot \left[\frac{\mathbf{v}}{v}\frac{\partial f_0}{\partial v} + \frac{\mathbf{f}_1}{v} + \frac{\mathbf{v}}{v}\frac{\partial}{\partial v}\left(\frac{\mathbf{f}_1}{v}\right) \cdot \mathbf{v}\right]. \tag{2.15}$$

The magnetic field term follows along similar lines:

$$(\mathbf{v} \times \mathbf{B}) \cdot \frac{\partial}{\partial \mathbf{v}} \left(f_0 + \mathbf{f}_1 \cdot \frac{\mathbf{v}}{v} \right) \;=\; (\mathbf{v} \times \mathbf{B}) \cdot \left[\frac{\mathbf{v}}{v} \frac{\partial f_0}{\partial v} + \frac{\mathbf{f}_1}{v} + \frac{\mathbf{v}}{v} \frac{\partial}{\partial v} \left(\frac{\mathbf{f}_1}{v} \right) \cdot \mathbf{v} \right] \quad (2.16)$$

$$\;=\; (\mathbf{B} \times \mathbf{f}_1) \cdot \frac{\mathbf{v}}{v}. \quad (2.17)$$

Finally, we consider the collision term. Looking at the Boltzmann equation as a continuity condition, S can be viewed as the outflow of electrons from a volume $d\mathbf{v}$ due to collisions:

$$S \, d\mathbf{v} = \int \rho_c \, d\Omega. \quad (2.18)$$

Here, ρ_c is the net electron velocity flux directed radially outward from $d\mathbf{v}$ due to collisions, and the integral is taken over all solid angles. The contributions to ρ_c consist of electrons scattering in and out of the volume $d\mathbf{v}$:

$$\rho_c = \rho_{out} - \rho_{in}. \quad (2.19)$$

Concerning scattering out of $d\mathbf{v}$, we can write the number of electrons scattered into an angle (θ', ϕ') per unit time, per unit volume, as

$$\rho_{out} = v f(\mathbf{v}) \, d\mathbf{v} \sum_j N_j I_j(v, \theta'). \quad (2.20)$$

The factor $v f(\mathbf{v}) \, d\mathbf{v}$ is the flux of electrons in configuration space, and the summation is the scattering cross section per unit volume, with N_j being the jth molecular species density, and I_j being the jth differential cross section, assumed to be azimuthally symmetric. We approximate the molecular scatterers as being infinitely massive, and therefore their dynamics do not play a role.

In a similar manner, electrons enter $d\mathbf{v}$ by scattering from some other location $d\mathbf{v}'$ through a scattering angle of (θ', ϕ'):

$$\rho_{in} = v' f(\mathbf{v}') \, d\mathbf{v}' \sum_j N_j I_j(v', \theta'). \quad (2.21)$$

We insert Equations (2.20) and (2.21) into (2.18) to arrive at the following collision integral:

$$S\,d\mathbf{v} = \sum_j N_j \int [v f(\mathbf{v}) I_j(v, \theta')\,d\mathbf{v} - v' f(\mathbf{v}') I_j(v', \theta')\,d\mathbf{v}']\,d\Omega'. \qquad (2.22)$$

If the collisions are completely elastic, then $v = v'$. This condition implies that the differential volumes $d\mathbf{v}$ and $d\mathbf{v}'$ are the same size. We can write

$$S = \sum_j N_j v \int [f(v, \theta, \phi) - f(v, \theta', \phi')] I_j(v, \theta')\,d\Omega'. \qquad (2.23)$$

Expanding f to first order yields:

$$S = \sum_j N_j v \int \left[f_0(v) + \mathbf{f}_1(v) \cdot \frac{\mathbf{v}}{v} - f_0(v) - \mathbf{f}_1(v) \cdot \frac{\mathbf{v}'}{v} \right] I_j(v, \theta')\,d\Omega' \qquad (2.24)$$

$$= \mathbf{f}_1 \cdot \sum_j N_j v \int \left(\frac{\mathbf{v}}{v} - \frac{\mathbf{v}'}{v} \right) I_j(v, \theta')\,d\Omega'. \qquad (2.25)$$

Since both $\mathbf{v}/v = \hat{\mathbf{z}}$ and I_j are azimuthally symmetric in the primed coordinates, only the azimuthally symmetric part of $\mathbf{v}'$ survives the integration, and we find

$$S = \frac{\mathbf{v}}{v} \cdot \mathbf{f}_1 \left[\sum_j N_j v \int (1 - \cos \theta') I_j(v, \theta')\,d\Omega' \right]. \qquad (2.26)$$

The quantity in brackets will be defined as ν:

$$S = \frac{\mathbf{v}}{v} \cdot \mathbf{f}_1 \nu. \qquad (2.27)$$

We now combine Equations (2.12), (2.15), (2.17), and (2.27), to arrive at the Boltzmann equation accurate to first order in the spherical harmonic expansion of f:

$$\frac{\partial f_0}{\partial t} - \frac{e}{m} \mathbf{E} \cdot \left[\frac{\mathbf{f}_1}{v} + \frac{\mathbf{v}}{v} \frac{\partial}{\partial v} \left(\frac{\mathbf{f}_1}{v} \right) \cdot \mathbf{v} \right] \qquad (2.28)$$

$$+ \frac{\mathbf{v}}{v} \cdot \left[\frac{\partial \mathbf{f}_1}{\partial t} - \frac{e}{m} \left(\mathbf{E} \frac{\partial f_0}{\partial v} + \mathbf{B} \times \mathbf{f}_1 \right) + \mathbf{f}_1 \nu \right] = 0. \qquad (2.29)$$

This is a scalar equation with four unknowns, namely the amplitudes (f_{000}, f_{110}, f_{111}, f_{100}) of the spherical harmonics to first order (Y_{000}, Y_{110}, Y_{111}, Y_{100}). The orthogonality property of the spherical harmonics will let us decompose Equation (2.29) into four scalar equations for the harmonic amplitudes.

2.1.2 Solution of the harmonic amplitudes

To solve for (f_{000}, f_{110}, f_{111}, f_{100}), we multiply Equation (2.29) by one of (Y_{000}, Y_{110}, Y_{111}, Y_{100}) and integrate over all angles. The outcomes of this process are moments of Equation (2.29), often referred to as the *spherical harmonic moment equations*, or more succinctly, the *moment equations*.

Beginning with the Y_{000} moment, we multiply Equation (2.29) by 1 and integrate over angle. The $(\mathbf{v}/v) \cdot [\ ]$ term on the second line of Equation (2.29) is of the form $(Y_{110}\hat{\mathbf{x}} + Y_{111}\hat{\mathbf{y}} + Y_{100}\hat{\mathbf{z}}) \cdot [\ ]$ and vanishes in the angle integration. Therefore we need only deal with the first line:

$$\int \left\{ \frac{\partial f_0}{\partial t} - \frac{e}{m}\mathbf{E} \cdot \left[\frac{\mathbf{f}_1}{v} + \frac{\mathbf{v}}{v}\frac{\partial}{\partial v}\left(\frac{\mathbf{f}_1}{v}\right) \cdot \mathbf{v} \right] \right\} d\Omega = 0. \tag{2.30}$$

The first two terms are trivial. We remove a factor of 4π and write the third term as a double contraction:

$$\frac{\partial f_0}{\partial t} - \frac{e}{m}\left[\mathbf{E} \cdot \frac{\mathbf{f}_1}{v} + \frac{v\mathbf{E}}{4\pi}\frac{\partial}{\partial v}\left(\frac{\mathbf{f}_1}{v}\right) : \int \frac{\mathbf{v}\mathbf{v}}{v^2}d\Omega \right] = 0. \tag{2.31}$$

The off-diagonal terms in the dyadic $\mathbf{v}\mathbf{v}/v^2$ involve odd powers of $\cos\phi$ or $\sin\phi$ and integrate to zero over the azimuthal coordinate. The diagonal terms integrate to $4\pi/3$. The double contraction of a tensor A with the identity tensor is the trace of A, thus

$$\frac{\partial f_0}{\partial t} - \frac{e\mathbf{E}}{m} \cdot \left[\frac{\mathbf{f}_1}{v} + \frac{v}{3}\frac{\partial}{\partial v}\left(\frac{\mathbf{f}_1}{v}\right) \right] = 0, \tag{2.32}$$

which can be written

$$\frac{\partial f_0}{\partial t} - \frac{e}{3mv^2}\mathbf{E} \cdot \frac{\partial}{\partial v}v^2\mathbf{f}_1 = 0. \tag{2.33}$$

This is the Y_{000} moment equation.

We now calculate the anisotropic Y_{110}, Y_{111}, and Y_{100} moments of Equation (2.29). Isotropic terms of Equation (2.29) will integrate to zero. This is the case for the first time derivative term. Recalling Equation (2.31), the $\mathbf{E} \cdot [\ ]$ term of Equation (2.29) involves an isotropic part $\mathbf{E} \cdot \mathbf{f}_1$, and an anisotropic part of the form

$$\int Y_{110} \frac{\mathbf{vv}}{v^2} \, d\Omega. \tag{2.34}$$

After writing out Y_{110} as $\sin \theta \cos \phi$ and likewise for the six distinct entries in the dyadic, one will find that all integrals over the sphere are zero. This same result holds for Y_{111} and Y_{100}. Therefore we need only consider the second line of Equation (2.29) for the three anisotropic moments. We start with Y_{110} and compute

$$\int \sin \theta \cos \phi \frac{\mathbf{v}}{v} \cdot \left[\frac{\partial \mathbf{f}_1}{\partial t} - \frac{e}{m} \left(\mathbf{E} \frac{\partial f_0}{\partial v} + \mathbf{B} \times \mathbf{f}_1 \right) + \mathbf{f}_1 \nu \right] d\Omega = 0. \tag{2.35}$$

Y_{110} is orthogonal to the $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ components of $\mathbf{v}$, thus we have

$$\left[\frac{\partial \mathbf{f}_1}{\partial t} - \frac{e}{m} \left(\mathbf{E} \frac{\partial f_0}{\partial v} + \mathbf{B} \times \mathbf{f}_1 \right) + \mathbf{f}_1 \nu \right]_x \int \sin^3 \theta \cos^2 \phi \, d\theta d\phi = 0, \tag{2.36}$$

where $[\ ]_x$ denotes the $\hat{\mathbf{x}}$ component, and the integral evaluates to $4\pi/3$. The Y_{111} and Y_{100} moments yield similar results for the $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ components. Rewriting these three scalar moment equations in vector form leaves us with

$$\frac{\partial \mathbf{f}_1}{\partial t} - \frac{e}{m} \left(\mathbf{E} \frac{\partial f_0}{\partial v} + \mathbf{B} \times \mathbf{f}_1 \right) + \mathbf{f}_1 \nu = 0. \tag{2.37}$$

This vector equation, along with the scalar equation (2.33) are a system of four scalar moment equations which can be used to solve for the variables (f_{000}, f_{100}, f_{110}, f_{111}). Henceforth we will refer to Equation (2.33) as the scalar moment equation, and Equation (2.37) as the vector moment equation.

2.1.3 Interpretation

We now look at the useful physical information contained in f_0 and $\mathbf{f}_1$ concerning bulk parameters such as density, current, and energy.

The first quantity of interest is the electron density:

$$n = \int f(\mathbf{v})\, d\mathbf{v} \tag{2.38}$$

$$= \sum_{lms} \int f_{lms}(v)v^2\, dv \int Y_{lms}(\theta, \phi)\, d\Omega. \tag{2.39}$$

All harmonics except Y_{000} will integrate to zero over angle, thus

$$n = \int f_{000}(v)v^2\, dv \int Y_{000}\, d\Omega \tag{2.40}$$

$$= 4\pi \int f_0 v^2\, dv. \tag{2.41}$$

Therefore electron number density can be recovered from f_0 alone by integrating over v, with the factor of $4\pi v^2$ to account for the size of the velocity volume element as one moves outward from the origin.

This result suggests that a conservation equation for the density can be derived from the scalar moment equation (2.33) by integrating over v:

$$4\pi \int \left(\frac{\partial f_0}{\partial t} - \frac{e}{3mv^2} \mathbf{E} \cdot \frac{\partial}{\partial v} v^2 \mathbf{f}_1 \right) v^2\, dv = 0. \tag{2.42}$$

Since $\mathbf{E}$ does not depend on v, we can pull it out of the integral:

$$\frac{\partial n}{\partial t} + \frac{4\pi e}{3m} \mathbf{E} \cdot \int \frac{\partial}{\partial v} v^2 \mathbf{f}_1\, dv = 0. \tag{2.43}$$

The quantity $v^2 \mathbf{f}_1$ vanishes at both limits, so we have simply

$$\frac{\partial n}{\partial t} = 0. \tag{2.44}$$

This result should not be surprising—it is the fluid equation of continuity for an assumed homogeneous plasma. While this example is somewhat trivial, it does begin

to illustrate the connection between the moment equations and the familiar fluid equations.

The next quantity to be investigated is the bulk electron motion $\overline{\mathbf{v}}$.

$$\overline{\mathbf{v}} = \frac{1}{n} \int \mathbf{v} f(\mathbf{v}) \, d\mathbf{v} \tag{2.45}$$

$$= \frac{1}{n} \sum_{lms} \int f_{lms}(v) v^2 \, dv \int \mathbf{v} Y_{lms}(\theta, \phi) \, d\Omega \tag{2.46}$$

$$= \frac{4\pi}{3n} \int \mathbf{f}_1 v^3 \, dv. \tag{2.47}$$

Therefore the bulk electron motion is determined entirely by $\mathbf{f}_1$. Because we are ignoring ion dynamics, the electric current is simply

$$\mathbf{J} = -en\overline{\mathbf{v}} \tag{2.48}$$

$$= -\frac{4\pi e}{3} \int \mathbf{f}_1 v^3 \, dv. \tag{2.49}$$

In the same manner as for the density, a conservation equation arises from integrating the vector moment equation (2.37) over v:

$$\frac{4\pi}{3} \int \left[\frac{\partial \mathbf{f}_1}{\partial t} - \frac{e}{m} \left(\mathbf{E} \frac{\partial f_0}{\partial v} + \mathbf{B} \times \mathbf{f}_1 \right) + \mathbf{f}_1 \nu \right] v^3 \, dv = 0. \tag{2.50}$$

The velocity derivative is integrated by parts, and we multiply by m to put this in a recognizable form:

$$\frac{\partial}{\partial t}(nm\overline{\mathbf{v}}) = -en(\mathbf{E} + \overline{\mathbf{v}} \times \mathbf{B}) - nm\nu_e \overline{\mathbf{v}}, \tag{2.51}$$

where ν_e is an effective collision frequency assumed to be given by

$$\overline{\mathbf{v}}\nu_e = \overline{\mathbf{v}\nu} \tag{2.52}$$

$$= -\frac{4\pi}{3n} \int \mathbf{f}_1 \nu v^3 \, dv. \tag{2.53}$$

Equation (2.51) is the fluid equation for conservation of momentum density. Thus

the vector moment equation (2.37) can be thought of as the v-dependent counterpart of the fluid momentum conservation equation.

Had we expanded f to second order, and derived a moment equation for $\mathbf{f}_2$, we might expect that it would give rise to a pressure tensor relation. Alas this analogy is not entirely correct as one will find that $\mathbf{f}_2$ corresponds to the $(\mathbf{v}\mathbf{v} - \frac{1}{3}v^2\mathbf{I}_2)$-moment of the Boltzmann equation, or in other words, the anisotropic component of $\overline{\mathbf{v}\mathbf{v}}$ [*Shkarofsky et al.*, 1966]. The isotropic part of $\overline{\mathbf{v}\mathbf{v}}$, namely the trace $\overline{\mathbf{v}\cdot\mathbf{v}}$, is contained in f_0. By multiplying $\overline{\mathbf{v}\cdot\mathbf{v}}$ by $m/2$, we can interpret this quantity as the average electron energy. The energy is a convenient scalar measure of the width of the distribution f_0 and some use of it will be made in later chapters. The moment can be calculated as follows:

$$\overline{\mathbf{v}\cdot\mathbf{v}} = \frac{1}{n}\int v^2 f(\mathbf{v})\,d\mathbf{v} \tag{2.54}$$

$$= \frac{1}{n}\sum_{lms}\int f_{lms}(v)v^4\,dv\int Y_{lms}(\theta,\phi)\,d\Omega \tag{2.55}$$

$$= \frac{4\pi}{n}\int f_0 v^4\,dv. \tag{2.56}$$

Thus the average electron energy arises from f_0 alone. A conservation equation for electron energy density arises from integrating the scalar moment equation (2.33) over v:

$$4\pi\frac{m}{2}\int\left(\frac{\partial f_0}{\partial t} - \frac{e}{3mv^2}\mathbf{E}\cdot\frac{\partial}{\partial v}v^2\mathbf{f}_1\right)v^4\,dv = 0. \tag{2.57}$$

We remove the electric field from the integral and integrate by parts:

$$\frac{\partial}{\partial t}\left(n\frac{m\overline{\mathbf{v}\cdot\mathbf{v}}}{2}\right) = -e\mathbf{E}\cdot\frac{4\pi}{3}\int \mathbf{f}_1 v^3\,dv. \tag{2.58}$$

The quantity $m\overline{\mathbf{v}\cdot\mathbf{v}}/2$ is the average electron energy ϵ. Thus

$$\frac{\partial}{\partial t}(n\epsilon) = \mathbf{E}\cdot\mathbf{J}. \tag{2.59}$$

This is the conservation equation for the total electron energy density. It is useful to

separate the energy into ordered and thermal components as follows:

$$\overline{\mathbf{v} \cdot \mathbf{v}} = \overline{\mathbf{v}} \cdot \overline{\mathbf{v}} - \overline{(\mathbf{v} - \overline{\mathbf{v}}) \cdot (\mathbf{v} - \overline{\mathbf{v}})}. \tag{2.60}$$

The ordered $\overline{\mathbf{v}} \cdot \overline{\mathbf{v}}$ component is computed by magnitude-squaring $\overline{\mathbf{v}}$ found earlier. The thermal component is then the difference of $\overline{\mathbf{v} \cdot \mathbf{v}}$ and $\overline{\mathbf{v}} \cdot \overline{\mathbf{v}}$.

The ordered and thermal components of the energy density have corresponding conservation equations. To arrive at a conservation equation for the ordered electron kinetic energy density, we take the dot product of $\overline{\mathbf{v}}$ and the momentum conservation equation (2.51):

$$\overline{\mathbf{v}} \cdot \frac{\partial}{\partial t}(nm\overline{\mathbf{v}}) = \overline{\mathbf{v}} \cdot [-en(\mathbf{E} + \overline{\mathbf{v}} \times \mathbf{B}) - nm\nu_e \overline{\mathbf{v}}] \tag{2.61}$$

$$\frac{\partial}{\partial t}\left(\frac{nm}{2}\overline{\mathbf{v}} \cdot \overline{\mathbf{v}}\right) = \mathbf{E} \cdot \mathbf{J} - nm\nu_e \overline{\mathbf{v}} \cdot \overline{\mathbf{v}}. \tag{2.62}$$

As suggested by Equation (2.60), a conservation equation for the internal heat kinetic energy density $n\epsilon_i$ is obtained by subtracting Equation (2.62) from (2.59):

$$\frac{\partial}{\partial t}(n\epsilon_i) = nm\nu_e \overline{\mathbf{v}} \cdot \overline{\mathbf{v}}. \tag{2.63}$$

This is a key result. According to our model, what is happening is that the electric field accelerates electrons in an ordered manner, but the electron direction of travel is quickly randomized through collisions with molecules. This random motion is what we characterize as internal heat kinetic energy. Equation (2.63) describes how this heat energy evolves with time: the electron population will continue to get hotter so long as the collision frequency and the electron velocity $\overline{\mathbf{v}}$ remain finite. This will be the case until the electric field is removed and $\overline{\mathbf{v}}$ goes to zero, at which point our simple model predicts that the electrons stay heated indefinitely. In practice this is unlikely, as we would expect loss processes to eventually cool the electron population. Energy loss channels at low electron energies include molecular recoil during elastic collisions, and excitation of rotational quantum levels of molecules.

2.2 Energy loss processes

Energy loss processes were neglected in the earlier treatment of the Boltzmann collision integral when it was approximated that the electron energy was unchanged during a collision with a molecular scatterer [Eq. (2.23)]. We now relax this approximation and treat the integral in a more rigorous fashion.

2.2.1 Molecular recoil

The first loss channel is the effect of molecular recoil during an electron-molecule elastic collision. In other words, some of the electron energy is transfered to the molecule, and the electron slows down. We recall the collision integral [Eq. (2.22)]:

$$S\,d\mathbf{v} = \sum_j N_j \int \left[vf(\mathbf{v})I_j(v,\theta')\,d\mathbf{v} - v'f(\mathbf{v}')I_j(v',\theta')\,d\mathbf{v}' \right] d\Omega'. \qquad (2.64)$$

Here, the first term represents the quantity of electrons scattered out of $d\mathbf{v}$ at an angle θ'. It does not matter how much energy they lose upon leaving $d\mathbf{v}$, but only that they have indeed departed.

The second term is quite different. In this term $\mathbf{v}'$ denotes the velocity-space location of electrons before they are scattered into the element $d\mathbf{v}$. Thus v' will be larger than v by the amount which the electron is slowed down when scattered at angle θ'. In order to establish a connection between v and v' we need to know how much electron energy is lost during an electron-molecule collision.

In analyzing energy loss, some notational confusion is possible. Equation (2.64) gives the number of particles scattered out of, and in to, element $d\mathbf{v}$. In the first term, the primed coordinates give the location of the electron after the collision, or in other words, the velocity after the collision is v'. In the second term, the primed coordinates give the location of the electron before the collision, so the velocity after the collision is v. Since v' appears only in the second term, it is logical to use v' to denote velocity before the collision in the following discussion of particle collision laws.

We consider the energy conservation relation for an electron-molecule collision with the molecule initially at rest:

$$\frac{1}{2}mv'^2 = \frac{1}{2}MV^2 + \frac{1}{2}mv^2. \qquad (2.65)$$

The lower case script refers to electrons and the upper case script refers to molecules. Solving for the velocity difference $v' - v$, we find the following approximate relationship:

$$v' - v \approx \frac{M}{2vm}\mathbf{V} \cdot \mathbf{V}. \qquad (2.66)$$

Using momentum conservation, we can eliminate $\mathbf{V}$:

$$
\begin{aligned}
v' - v &= \frac{m}{2vM}(\mathbf{v}' - \mathbf{v}) \cdot (\mathbf{v}' - \mathbf{v}) & (2.67) \\
&\approx \frac{m}{vM}(v'^2 - vv'\cos\theta') & (2.68) \\
&\approx \frac{mv'}{M}(1 - \cos\theta'). & (2.69)
\end{aligned}
$$

Thus the change in velocity scales with the pre-collision velocity and depends strongly on the scattering angle θ'. The angular dependence is as one might expect—the amount of energy exchanged with the molecule increases monotonically with scattering angle.

We are now in a position to attack the collision integral. The first order of business is to relate the $d\mathbf{v}$ and $d\mathbf{v}'$ factors. We have

$$
\begin{aligned}
d\mathbf{v}' &= v'^2\, dv'\, d\Omega' & (2.70) \\
&= v'^2 \frac{dv'}{dv}\, dv\, d\Omega & (2.71) \\
&= \frac{v'^2}{v^2}\frac{dv'}{dv}\, d\mathbf{v}. & (2.72)
\end{aligned}
$$

Thus we need dv'/dv. From Equation (2.69) we find that

$$dv' = \frac{v'}{v}\, dv. \qquad (2.73)$$

Combining Equations (2.72) and (2.73) yields

$$dv' = \frac{v'^3}{v^3} dv. \tag{2.74}$$

The volume element can now be removed from the collision integral, leaving us with

$$S = \sum_j N_j v \int \left[f(\mathbf{v}) I_j(v, \theta') - \frac{v'^4}{v^4} f(\mathbf{v}') I_j(v', \theta') \right] d\Omega'. \tag{2.75}$$

The second term in brackets in Equation (2.64) can be expanded in Taylor series around v since $|v' - v| \ll v, v'$:

$$S = \sum_j N_j v \left\{ \int [f(v, \theta) - f(v, \theta')] I_j(v, \theta') \, d\Omega' - \int \frac{v' - v}{v^4} \frac{\partial}{\partial v} [v^4 f(v, \theta') I_j(v, \theta')] \, d\Omega' \right\}. \tag{2.76}$$

The first integral was treated in Equation (2.23) and it was shown to contribute to the vector moment equation for $\mathbf{f}_1$. We are interested, however, in modifications to the scalar moment equation, thus we ignore the first term and consider only the isotropic part of the second term (with $f \approx f_0$), which we will denote as S_0:

$$S_0 = -\sum_j \frac{N_j}{v^3} \int (v' - v) \frac{\partial}{\partial v} [v^4 f_0(v) I_j(v, \theta')] \, d\Omega'. \tag{2.77}$$

We insert Equation (2.69), and move the sum and integrals into the derivative:

$$S_{0e} \approx -\frac{m}{Mv^2} \frac{\partial}{\partial v} \left\{ v^3 f_0 \left[\sum_j N_j v \int (1 - \cos \theta') I_j(v, \theta') \, d\Omega' \right] \right\} \tag{2.78}$$

$$= -\frac{m}{Mv^2} \frac{\partial}{\partial v} \left(v^3 f_0 \nu \right) \tag{2.79}$$

$$= -\frac{mN}{Mv^2} \frac{\partial}{\partial v} \left(v^4 f_0 \sigma_t \right) \tag{2.80}$$

The $0e$ subscript denotes the elastic part of S_0. Here we have used the definition of ν from Equation (2.27), and converted to total cross section using $\nu = N\sigma_t v$. N and σ_t are density and cross section averaged over molecular species, in this case a 4:1 nitrogen/oxygen ratio, which we will denote as "air". We recall that ν was associated
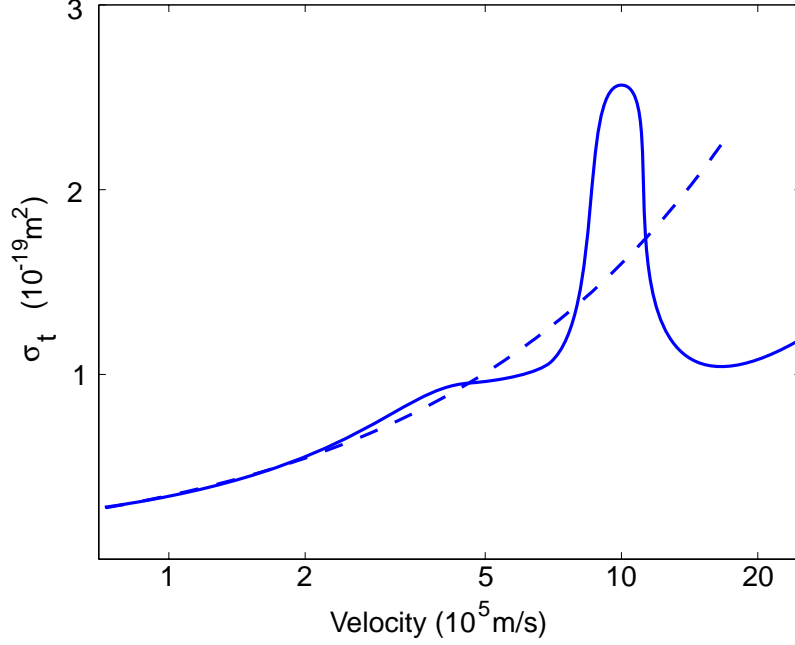
Figure 2-1: Cross section σ_t for electron-air molecular collisions.

with momentum relaxation [Eq. (2.37)], and thus σ_t is referred to as the cross section for "momentum transfer". However, we have shown here that when molecular recoil is considered, this cross section is also relevant to energy loss associated with f_0. The momentum transfer cross section σ_t is calculated [*Gurevich*, 1978] using experimental data [*Crompton and Sutton*, 1952, *Engelhardt et al.*, 1964, *Hake and Phelps*, 1967] and is given in Figure 2-1. In the figure, the dashed curve

$$\sigma_t = Cv^{2/3}, \qquad (2.81)$$

where $C = 1.6 \times 10^{-23}$ m$^{4/3}$s$^{2/3}$, is a good approximation to σ_t for electron velocities less than 8×10^5 m/s. The large peak near 8×10^5 m/s is due to the influence of inelastic nitrogen vibrational excitation on the momentum transfer cross section.

2.2.2 Rotational excitation

We now consider inelastic collisions of electrons with molecules. Elastic collisions tend to dominate σ_t (at least in the range $v < 8 \times 10^5$ m/s), so the effect of inelastic

collisions on $\mathbf{f}_1$ will be ignored. Conversely, we can ignore the effect of $\mathbf{f}_1$ on the inelastic collision theory given below if we assume weak anisotropy ($|\mathbf{f}_1| \ll |f_0|$).

In a rotational excitation, the molecule gains or loses angular momentum via the exchange of energy quanta with the electrons. This energy exchange is different from the elastic case because the electrons can either lose *or* gain energy. Collisions resulting in electron energy loss are referred to as "collisions of the first kind", whereas those resulting in energy gain are "collisions of the second kind".

As before, we evaluate the contributions to ρ_c, the net electron velocity flux out of $d\mathbf{v}$. The first outgoing contribution concerns electrons which begin with energy $\epsilon = (m/2)v^2$, excite the molecule from state k to the more energetic state $k + j$, and finish with energy $\epsilon - \epsilon_{kj} = (m/2)v'^2$:

$$\rho_{out,1} = v f_0(v)\, d\mathbf{v} \sum_{k,j} N_k I_{kj}(v, \theta'). \tag{2.82}$$

The "1" subscript denotes a collision of the first kind. N_k is the density of molecules in the kth state. To simplify the notation, we enumerate the k (rotational energy) states of a single molecular species and omit the summation over species. The final loss rates should be averaged over the molecular composition of air, as was done for molecular recoil.

In a manner similar to the above, a collision of the second kind would involve the electron encountering a molecule which de-excites from state k to the less energetic state $k - j$, leaving the electron with an energy $\epsilon + \epsilon_{k,-j} = (m/2)v'^2$:

$$\rho_{out,2} = v f_0(v)\, d\mathbf{v} \sum_{k,j} N_k I_{k,-j}(v, \theta'). \tag{2.83}$$

The arrival of electrons follows in the same manner. In a collision of the first kind, electrons begin with energy $\epsilon + \epsilon_{kj} = (m/2)v'^2$ and finish with $\epsilon = (m/2)v^2$:

$$\rho_{in,1} = v' f_0(v')\, d\mathbf{v}' \sum_{k,j} N_k I_{kj}(v', \theta'). \tag{2.84}$$

The last case involves electrons beginning with energy $\epsilon - \epsilon_{k,-j} = (m/2)v'^2$ and

increasing to energy $\epsilon = (m/2)v^2$:

$$\rho_{in,2} = v' f_0(v') \, d\mathbf{v}' \sum_{k,j} N_k I_{k,-j}(v', \theta'). \tag{2.85}$$

The net outgoing electron flux is thus:

$$\rho_c = \rho_{out,1} + \rho_{out,2} - \rho_{in,1} - \rho_{in,2}. \tag{2.86}$$

Before we attempt to combine these expressions, a little simplification is in order. The $d\mathbf{v}'$ factor can be related to $d\mathbf{v}$ by examining the energy relation:

$$\frac{m}{2} v'^2 = \frac{m}{2} v^2 \pm \epsilon_{k,\pm j}. \tag{2.87}$$

Differentiating, we find that

$$dv' = \frac{v}{v'} \, dv. \tag{2.88}$$

Note that this is the reciprocal of the relation for elastic collisions [Eq. (2.73)]. From Equation (2.72) it follows that

$$d\mathbf{v}' = \frac{v'}{v} \, d\mathbf{v}. \tag{2.89}$$

We insert this volume element into Equations (2.82) through (2.85), and integrate over angle:

$$S_0 = \sum_{k,j} N_k \Bigg[v f_0(v) \sigma_{kj}(v) + v f_0(v) \sigma_{k,-j}(v) \tag{2.90}$$

$$-\frac{v'^2}{v} f_0(v') \sigma_{kj}(v') \Big|_{v'=\sqrt{2(\epsilon+\epsilon_{kj})/m}} - \frac{v'^2}{v} f_0(v') \sigma_{k,-j}(v') \Big|_{v'=\sqrt{2(\epsilon-\epsilon_{k,-j})/m}} \Bigg]. \tag{2.91}$$

Here, $\sigma(v)$ is the angle-integrated (total) cross section. At this point, it is useful to move to energy variables. Equation (2.91) is:

$$S_0 = \frac{2}{mv} \sum_{k,j} N_k \Bigg[\epsilon f_0(\epsilon) \sigma_{kj}(\epsilon) + \epsilon f_0(\epsilon) \sigma_{k,-j}(\epsilon) \tag{2.92}$$

$$-\epsilon' f_0(\epsilon') \sigma_{kj}(\epsilon') \Big|_{\epsilon'=\epsilon+\epsilon_{kj}} - \epsilon' f_0(\epsilon') \sigma_{k,-j}(\epsilon') \Big|_{\epsilon'=\epsilon-\epsilon_{k,-j}} \Bigg]. \tag{2.93}$$

45

We expand to first order in Taylor series, and the zero-order term drops out:

$$S_0 = -\frac{2}{mv} \sum_{k,j} N_k \frac{\partial}{\partial \epsilon} \left[\epsilon f_0 (\epsilon_{kj} \sigma_{kj} - \epsilon_{k,-j} \sigma_{k,-j}) \right].$$ (2.94)

To evaluate this expression, we will borrow the cross sections from the published literature [*Gerjouy and Stein*, 1955]:

$$\sigma_{k,k+2}(\epsilon) = \sigma_r \frac{(k+2)(k+1)}{(2k+3)(2k+1)} \left[1 - \frac{2(2k+3)B_0}{\epsilon} \right]^{1/2}$$ (2.95)

$$\sigma_{k,k-2}(\epsilon) = \sigma_r \frac{k(k-1)}{(2k-1)(2k+1)} \left[1 + \frac{2(2k-1)B_0}{\epsilon} \right]^{1/2}$$ (2.96)

$$\sigma_r = \frac{8\pi Q^2 a_0^2}{15}.$$ (2.97)

The cross sections are nonzero for state changes of k to $k \pm 2$, and otherwise zero. B_0 and Q are constants, which for nitrogen and oxygen are given by [*Herzberg*, 1950]:

$$B_0(N_2) = 2.48 \cdot 10^{-4} \, \text{eV}$$ (2.98)

$$B_0(O_2) = 1.79 \cdot 10^{-4} \, \text{eV}$$ (2.99)

$$Q(N_2) = 1.04$$ (2.100)

$$Q(O_2) = 1.8.$$ (2.101)

a_0 is the Bohr radius

$$a_0 = 5.29 \times 10^{-11} \, \text{m},$$ (2.102)

and the energy levels are given by

$$\epsilon_k = B_0 k(k+1),$$ (2.103)

so that

$$\epsilon_{k,k+2} = \epsilon_{k+2} - \epsilon_k$$ (2.104)

$$= (4k+6)B_0$$ (2.105)

46

$$\epsilon_{k,k-2} = \epsilon_k - \epsilon_{k-2} \tag{2.106}$$

$$= (4k - 2)B_0. \tag{2.107}$$

In view of the fact that B_0 is tiny compared to typical electron energies ϵ in the plasma, we can ignore the bracketed factors in the rotational cross section expressions [Eqs. (2.95) and (2.96)], and insert the above results into Equation (2.94):

$$
\begin{aligned}
S_0 &= -\frac{2}{mv} \sum_k N_k \frac{\partial}{\partial \epsilon} \left\{ \epsilon f_0 B_0 \sigma_r \left[\frac{(4k+6)(k+2)(k+1)}{(2k+3)(2k+1)} - \frac{(4k-2)k(k-1)}{(2k-1)(2k+1)} \right] \right\} \\
&= -\frac{8 B_0 \sigma_r}{mv} \frac{\partial}{\partial \epsilon} (\epsilon f_0) \sum_k N_k. \tag{2.108}
\end{aligned}
$$

The summation over states is just the molecular density N. Going back to velocities, we find that

$$S_{0r} = -\frac{4 B_0 \sigma_r N}{mv^2} \frac{\partial}{\partial v} (v^2 f_0), \tag{2.109}$$

where $0r$ denotes S_0 due to rotational excitation.

2.2.3 Interpretation

The collision integral S_0 due to molecular recoil and rotational excitation has the compact form

$$S_0 = S_{0e} + S_{0r} \tag{2.110}$$

$$= -\frac{1}{2v^2} \frac{\partial}{\partial v} (R v^3 f_0), \tag{2.111}$$

where R is a sum of loss rates for the two processes:

$$R = R_e + R_r \tag{2.112}$$

$$R_e = \frac{2mNv\sigma_t}{M} \tag{2.113}$$

$$R_r = \frac{8 B_0 \sigma_r N}{mv}. \tag{2.114}$$

The rationale for defining an effective loss rate R in Equation (2.111) is clear when we add a loss term S_0 to the scalar moment equation (2.33). Ignoring the electric field, it is:

$$\frac{\partial f_0}{\partial t} - \frac{1}{2v^2}\frac{\partial}{\partial v}(Rv^3 f_0) = 0. \tag{2.115}$$

If we assume, for illustrative example, that R is constant over the region of interest, then we can easily calculate an energy density conservation relation [see Eq. (2.57)] by taking the $2\pi m v^4$-moment of Equation (2.115):

$$\frac{\partial}{\partial t}(n\epsilon) - \pi m R \int v^2 \frac{\partial}{\partial v}(v^3 f_0)\, dv = 0, \tag{2.116}$$

where $n\epsilon$ is the total electron energy density (equal to the internal heat kinetic energy density since there is no electric field to drive ordered motion). After integrating the second term by parts, we find that

$$\frac{\partial}{\partial t}(n\epsilon) + Rn\epsilon = 0. \tag{2.117}$$

Therefore the energy relaxation is exactly that of a first order system, decaying exponentially with a time constant $\tau = 1/R$, hence the rationale for defining R as a "loss rate". In the general case, of course, R has velocity dependence. However, the essential character of the energy relation is the same.

2.2.4 Effect of molecular motion

We have so far assumed that the molecules are stationary at the onset of a collision. This does not appear to be a problem since the ratio of velocities of electrons and molecules is on the order of $\sqrt{M/m} \approx 230$. Nevertheless, the molecular motion is important as we would expect the electrons to relax to a Maxwellian distribution with the temperature T of the molecules, as opposed to zero energy as predicted by Equation (2.117).

The loss rate defined in Equation (2.111) needs to be modified to account for the molecular velocity distribution. The main effect of the molecular velocity spread will

be to provide a small perturbation to the precollision electron velocity relative to the molecule:

$$\mathbf{v} = \mathbf{v}|_{T=0} + \delta\mathbf{v}, \tag{2.118}$$

which produces a perturbation to the distribution function:

$$f_0(\mathbf{v}) = f_0(\mathbf{v})|_{T=0} + \delta\mathbf{v}\frac{\partial f_0}{\partial v}. \tag{2.119}$$

We enter this correction into Equation (2.111):

$$S_0 = -\frac{1}{2v^2}\frac{\partial}{\partial v}\left[Rv^3\left(f_0 + \delta\mathbf{v}\frac{\partial f_0}{\partial v}\right)\right]. \tag{2.120}$$

We need to determine $\delta\mathbf{v}$ such that this expression produces the correct loss rate. The simplest route is the case of equilibrium, with no fields, such that $\frac{\partial f_0}{\partial t}, S_0 = 0$, and the electrons are Maxwell-distributed at the molecular temperature T. In this case, Equation (2.120) implies

$$0 = -\frac{1}{2v^2}\frac{\partial}{\partial v}\left[Rv^3\left(1 + \delta\mathbf{v}\frac{\partial}{\partial v}\right)n\left(\frac{m}{2\pi\kappa T}\right)^{3/2}\exp\left(-\frac{mv^2}{2\kappa T}\right)\right]. \tag{2.121}$$

Thus we must have

$$\delta\mathbf{v} = \frac{\kappa T}{mv}. \tag{2.122}$$

The corrected form of the collision integral is then [*Shkarofsky et al.*, 1966]:

$$S_0 = -\frac{1}{2v^2}\frac{\partial}{\partial v}\left[Rv^3\left(f_0 + \frac{\kappa T}{mv}\frac{\partial f_0}{\partial v}\right)\right]. \tag{2.123}$$

We hereby add it to the scalar moment equation (2.33) so that it now reads

$$\frac{\partial f_0}{\partial t} - \frac{e}{3mv^2}\mathbf{E}\cdot\frac{\partial}{\partial v}v^2\mathbf{f}_1 - \frac{1}{2v^2}\frac{\partial}{\partial v}\left[Rv^3\left(f_0 + \frac{\kappa T}{mv}\frac{\partial f_0}{\partial v}\right)\right] = 0. \tag{2.124}$$

2.3 Fourier series expansion

2.3.1 Expansion to first order

We now expand all quantities in Fourier series of the harmonics of an applied HF heater wave field, and retain zero- and first-harmonic terms. The first quantity to expand is the electric field itself. We need to account for two sources—an applied heater wave field $\mathbf{E}^1$ at a carrier frequency ω, and the constant electric field $\mathbf{E}^0$ due to the solar wind dynamo:

$$\mathbf{E}(t) = \mathbf{E}^0 + \mathbf{E}^1 e^{-i\omega t}. \tag{2.125}$$

The superscripts represent temporal harmonics, and the subscripts will continue to represent spherical harmonics. As per convention, we take the real part of the complex electric field to represent the physical electric field.

The next quantity to expand is the distribution function. All spherical harmonics will respond at all temporal harmonics of ω:

$$f_{lms}(t) = \sum_{q=0}^{\infty} f_{lms}^q e^{-iq\omega t}. \tag{2.126}$$

Of course, we have been focusing our attention on $f_0 = f_{000}$ and $\mathbf{f}_1 = f_{110}\hat{\mathbf{x}} + f_{111}\hat{\mathbf{y}} + f_{100}\hat{\mathbf{z}}$ in our development of the scalar and vector moment equations. These equations are coupled, nonlinear relations which lead to various feedback channels for the excitation of f_0 and $\mathbf{f}_1$. The analysis can be greatly simplified by identifying which channels are dominant. We start by considering the nonlinearities of the scalar and vector moment equations, respectively:

$$f_0 \quad \sim \quad \mathbf{E} \cdot \mathbf{f}_1 \tag{2.127}$$

$$\mathbf{f}_1 \quad \sim \quad \mathbf{E} f_0. \tag{2.128}$$

Therefore f_0^0, the zeroth harmonic of f_0, could be excited by the beating of $\mathbf{E}^1$ and $\mathbf{f}_1^1$. Alternatively, it could be excited by the beating of $\mathbf{E}^0$ and $\mathbf{f}_1^0$. Either way, once f_0^0 is excited, it can then beat with $\mathbf{E}^1$ to feed back into $\mathbf{f}_1^1$ or beat with $\mathbf{E}^0$ to feed

50

back into $\mathbf{f}_1^0$. These channels can be illustrated as follows:

$$
\begin{array}{ccc}
\boxed{f_0^0} & \xleftrightarrow{\;\mathbf{E}^0\;} & \boxed{\mathbf{f}_1^0} \\[4pt]
\mathbf{E}^1 \updownarrow & & \mathbf{E}^1 \updownarrow \\[4pt]
\boxed{\mathbf{f}_1^1} & \xleftrightarrow{\;\mathbf{E}^0\;} & \boxed{f_0^1} \\[4pt]
\mathbf{E}^1 \updownarrow & & \mathbf{E}^1 \updownarrow \\[4pt]
\boxed{f_0^2} & \xleftrightarrow{\;\mathbf{E}^0\;} & \boxed{\mathbf{f}_1^2} \\[4pt]
\vdots & & \vdots
\end{array}
$$
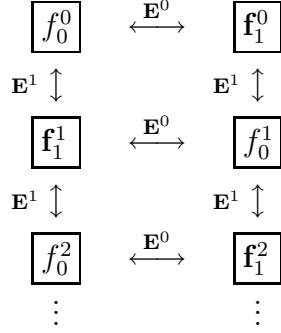
We now determine the dominant channels. First, $\mathbf{E}^0$ is on the order of millivolts, whereas $\mathbf{E}^1$ is on the order of volts, so we expect that the two vertical paths are essentially decoupled. Second, for most interactions we expect the ordering $f_0^0 \ll f_0^1 \ll f_0^2$. Thus the feedback between f_0^0 and $\mathbf{f}_1^1$ is almost certainly the dominant channel during plasma heating.

The component $\mathbf{f}_1^0$ is also of interest, as it will be shown shortly that it describes the antenna current in the ELF/VLF generation experiments. It has been shown in experiments [*Kasputin*, 1977] that this antenna current is linearly proportional to $\mathbf{E}^0$, which suggests that it arises from f_0^0 rather than via feedback between f_0^1 and $\mathbf{f}_1^0$. Thus the dominant channels are:

$$
\boxed{\mathbf{f}_1^1} \xleftrightarrow{\;\mathbf{E}^1\;} \boxed{f_0^0} \xrightarrow{\;\mathbf{E}^0\;} \boxed{\mathbf{f}_1^0}
$$
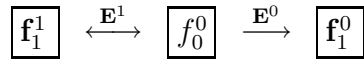
With this picture in mind, we expand f according to

$$
f = f_0^0 + (\mathbf{f}_1^0 + \mathbf{f}_1^1 e^{-i\omega t}) \cdot (\mathbf{v}/v). \tag{2.129}
$$

Inserting these expansions into the moment equations yields

$$
\begin{aligned}
\frac{\partial f_0^0}{\partial t} - \frac{e}{3mv^2} &\left(\mathbf{E}^0 + \mathbf{E}^1 e^{-i\omega t} \right) \cdot \frac{\partial}{\partial v} \left[v^2 (\mathbf{f}_1^0 + \mathbf{f}_1^1 e^{-i\omega t}) \right] \\
&- \frac{1}{2v^2} \frac{\partial}{\partial v} \left[R v^3 \left(f_0^0 + \frac{\kappa T}{mv} \frac{\partial f_0^0}{\partial v} \right) \right] = 0
\end{aligned} \tag{2.130}
$$

$$\frac{\partial}{\partial t}(\mathbf{f}_1^0 + \mathbf{f}_1^1 e^{-i\omega t}) - \frac{e}{m}\left[\left(\mathbf{E}^0 + \mathbf{E}^1 e^{-i\omega t}\right)\frac{\partial f_0^0}{\partial v} + \mathbf{B} \times (\mathbf{f}_1^0 + \mathbf{f}_1^1 e^{-i\omega t})\right]$$
$$+\nu(\mathbf{f}_1^0 + \mathbf{f}_1^1 e^{-i\omega t}) = 0. \qquad (2.131)$$

2.3.2   Solution of the harmonic amplitudes

We now determine the harmonic amplitudes f_0^0, $\mathbf{f}_1^0$, and $\mathbf{f}_1^1$ by separating the tempo-rally orthogonal parts of Equations (2.130) and (2.131). The formal process for doing so is to multiply the equations by $e^{iq\omega t}$ ($q = 0, 1, ...$), integrate over $0 < t < 2\pi/\omega$, and divide by $2\pi/\omega$. We need to be careful in multiplying complex quantities, since we desire to represent the product of their real parts, rather than the real part of their products. For example, the real (desired) part of the zeroth harmonic beat of $\mathbf{E}^1$ and $\mathbf{f}_1^1$ is given by

$$\left\langle \Re(\mathbf{E}^1 e^{-i\omega t}) \cdot \Re(\mathbf{f}_1^1 e^{-i\omega t})\right\rangle = \frac{1}{2}\Re(\mathbf{E}^{1*} \cdot \mathbf{f}_1^1). \qquad (2.132)$$

We integrate Equation (2.130) over a period with $q = 0$ and obtain for f_0^0:

$$\frac{\partial f_0^0}{\partial t} - \frac{e}{6mv^2}\frac{\partial}{\partial v}\left[v^2\Re(\mathbf{E}^{1*} \cdot \mathbf{f}_1^1)\right] - \frac{1}{2v^2}\frac{\partial}{\partial v}\left[Rv^3\left(f_0^0 + \frac{\kappa T}{mv}\frac{\partial f_0^0}{\partial v}\right)\right] = 0. \qquad (2.133)$$

Here we have ignored the product of $\mathbf{E}^0$ and $\mathbf{f}_1^0$, and have moved $\mathbf{E}^1$ inside the deriva-tive. Similarly, to get equations for $\mathbf{f}_0^1$ and $\mathbf{f}_1^1$, we integrate Equation (2.131) with $q = 0, 1$:

$$\mathbf{f}_1^q = -\mathbf{u}^q\frac{\partial f_0^0}{\partial v}, \qquad (2.134)$$

where $\mathbf{u}^q$ satisfies

$$\frac{d\mathbf{u}^q}{dt} + \nu\mathbf{u}^q = -\frac{e}{m}(\mathbf{E}^q + \mathbf{u}^q \times \mathbf{B}). \qquad (2.135)$$

Thus $\mathbf{f}_1$ follows the trajectory of an electron in the fields $\mathbf{E}$ and $\mathbf{B}$. If we assume that $\mathbf{B}$ is a constant field in the z-direction, and all quantities in Equation (2.135) vary with $e^{-iq\omega t}$, $q = 0, 1$, then

$$(\nu - iq\omega)\mathbf{u}^q + \omega_c\mathbf{u}^q \times \hat{\mathbf{z}} = -\frac{e}{m}\mathbf{E}^q. \qquad (2.136)$$

52

Solving for $\mathbf{u}^q$ and inserting into Equation (2.134) gives us the well-known results [*Shkarofsky et al.*, 1966]:

$$\mathbf{f}_1^0 = - \begin{bmatrix} \dfrac{\nu}{\omega_c^2 + \nu^2} & -\dfrac{\omega_c}{\omega_c^2 + \nu^2} & 0 \\[2ex] \dfrac{\omega_c}{\omega_c^2 + \nu^2} & \dfrac{\nu}{\omega_c^2 + \nu^2} & 0 \\[2ex] 0 & 0 & \dfrac{1}{\nu} \end{bmatrix} \left(\dfrac{e\mathbf{E}^0}{m} \dfrac{\partial f_0^0}{\partial v} \right) \qquad (2.137)$$

$$\mathbf{f}_1^1 = - \begin{bmatrix} \dfrac{\nu - i\omega}{\omega_c^2 + (\nu - i\omega)^2} & -\dfrac{\omega_c}{\omega_c^2 + (\nu - i\omega)^2} & 0 \\[2ex] \dfrac{\omega_c}{\omega_c^2 + (\nu - i\omega)^2} & \dfrac{\nu - i\omega}{\omega_c^2 + (\nu - i\omega)^2} & 0 \\[2ex] 0 & 0 & \dfrac{1}{\nu - i\omega} \end{bmatrix} \left(\dfrac{e\mathbf{E}^1}{m} \dfrac{\partial f_0^0}{\partial v} \right). \qquad (2.138)$$

Equations (2.133), (2.137), and (2.138) are the equations for the harmonic amplitudes f_0^0, $\mathbf{f}_1^0$, and $\mathbf{f}_1^1$.

2.3.3 Interpretation

The physically useful information contained in the above equations are the energy (associated with f_0^0) and the electric current (associated with $\mathbf{f}_1^q$). To properly interpret the energy, we need to choose a particular form for $\mathbf{E}^1$ in Equation (2.133). We consider the case of injection into the plasma along the magnetic field lines. In the lower ionosphere, in particular, the field polarization is not significantly affected by the plasma, so the field can be written:

$$\mathbf{E}^1 e^{-i\omega t} = \frac{E^1}{\sqrt{2}} a(t)(\hat{\mathbf{x}} \pm i\hat{\mathbf{y}})e^{-i\omega t}, \qquad (2.139)$$

where $\pm$ denotes the R- and L-mode cases, and $a(t)$ is a slowly varying (compared to $e^{-i\omega t}$) amplitude modulation of the carrier. The coordinate system here has $\hat{\mathbf{z}}$ pointing vertically downward (along the earth's magnetic field $\mathbf{B}$), $\hat{\mathbf{x}}$ pointing north, and $\hat{\mathbf{y}}$ pointing east. The factor of $\sqrt{2}$ ensures that the vertical power flux is $(E^1)^2 p(t)/2\eta$, where $p(t) = |a(t)|^2$ is a slowly-varying power modulation function and η is the impedance of free space (since the HF frequency is far above the electron plasma

frequency). We combine Equations (2.133), (2.138) and (2.139) to arrive at:

$$\frac{\partial f_0^0}{\partial t} = \frac{e^2 (E^1)^2 p(t)}{6m^2 v^2} \frac{\partial}{\partial v} \left[\frac{v^2 \nu}{(\omega \mp \omega_c)^2 + \nu^2} \frac{\partial f_0^0}{\partial v} \right] + \frac{1}{2v^2} \frac{\partial}{\partial v} \left[Rv^3 \left(f_0^0 + \frac{\kappa T}{mv} \frac{\partial f_0^0}{\partial v} \right) \right],$$
(2.140)

By taking the second moment of this relation [see Eqs. (2.56) and (2.117)] we arrive at an equation for the electron energy:

$$\frac{\partial \epsilon}{\partial t} = Q - L.$$
(2.141)

Thus the Boltzmann equation for f_0^0 [Eq. (2.140)] has the basic structure of a first-order heat equation, where the change in electron energy is the difference of heating and cooling terms. We will keep this basic structure in mind in later chapters.

With regards to $\mathbf{f}_1^0$ and $\mathbf{f}_1^1$, they can be interpreted as related to current:

$$\mathbf{J}^0 = -\frac{4\pi e}{3} \int \mathbf{f}_1^0 v^3 \, dv$$
(2.142)

$$\mathbf{J}^1 = -\frac{4\pi e}{3} \int \mathbf{f}_1^1 v^3 \, dv.$$
(2.143)

$\mathbf{J}^0$ is a current which varies slowly in response to the power modulation $p(t)$. It is the "antenna current" which produces ELF/VLF radiation, and is of great importance in Chapters 3 and 4. $\mathbf{J}^1$ is the quiver current, which is the simple oscillatory motion of electrons in the HF heater electric field, and is of lesser importance. Henceforth we will use the unsubscripted symbol $\mathbf{J}$ to denote the antenna current $\mathbf{J}^0$.

As a final point, it should be noted that $\mathbf{E}^0$ is not known *a priori* and it is therefore useful to talk about the ionospheric conductivity rather than the actual current $\mathbf{J}$. To recast the problem in terms of conductivity we insert Equation (2.137) into (2.142):

$$\mathbf{J} = \boldsymbol{\sigma} \mathbf{E}^0,$$
(2.144)

where the conductivity tensor is given by:

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_P & -\sigma_H & 0 \\ \sigma_H & \sigma_P & 0 \\ 0 & 0 & \sigma_0 \end{bmatrix}, \tag{2.145}$$

with

$$\sigma_P = -\frac{4\pi e^2}{3m} \int v^3 \frac{\nu}{\omega_c^2 + \nu^2} \frac{\partial f_0^0}{\partial v} \, dv \tag{2.146}$$

$$\sigma_H = -\frac{4\pi e^2}{3m} \int v^3 \frac{\omega_c}{\omega_c^2 + \nu^2} \frac{\partial f_0^0}{\partial v} \, dv \tag{2.147}$$

$$\sigma_0 = -\frac{4\pi e^2}{3m} \int v^3 \frac{1}{\nu} \frac{\partial f_0^0}{\partial v} \, dv. \tag{2.148}$$

The three conductivities are referred to as Pedersen, Hall, and specific. This is the low-frequency, homogeneous version of the conductivity tensor given by *Allis* [1956]. The conductivity tensor plays a central role in the antenna current models presented in Chapter 3.

Chapter 3

The Electrojet Antenna

The purpose of this chapter is to solve Equation (2.140),

$$\frac{\partial f_0^0}{\partial t} = \frac{e^2 (E^1)^2 p(t)}{6m^2 v^2} \frac{\partial}{\partial v} \left[\frac{v^2 \nu}{(\omega \mp \omega_c)^2 + \nu^2} \frac{\partial f_0^0}{\partial v} \right] + \frac{1}{2v^2} \frac{\partial}{\partial v} \left[R v^3 \left(f_0^0 + \frac{\kappa T}{mv} \frac{\partial f_0^0}{\partial v} \right) \right], \quad (3.1)$$

and from it determine the ELF/VLF antenna current.

3.1 The distribution function

Equation (3.1) will be solved in the limits where $p(t)$ varies either slowly or quickly compared to the characteristic thermal response time τ of f_0^0.

3.1.1 Slow modulation limit

The first case is where $p(t)$ varies slowly compared to τ such that we can assume $\frac{\partial f_0^0}{\partial t} = 0$. With the time derivative of Equation (3.1) set to zero, we have

$$0 = \frac{\partial}{\partial v} \left\{ \frac{e^2 (E^1)^2 p(t) v^2 \nu}{3m^2 [(\omega \mp \omega_c)^2 + \nu^2]} \frac{\partial f_0^0}{\partial v} + v^2 R \frac{\kappa T}{m} \frac{\partial f_0^0}{\partial v} + v^3 R f_0^0 \right\}. \quad (3.2)$$

The expression in braces must be a constant. To determine this constant, we consider the limit as v approaches zero and see that this constant is zero. Therefore we can
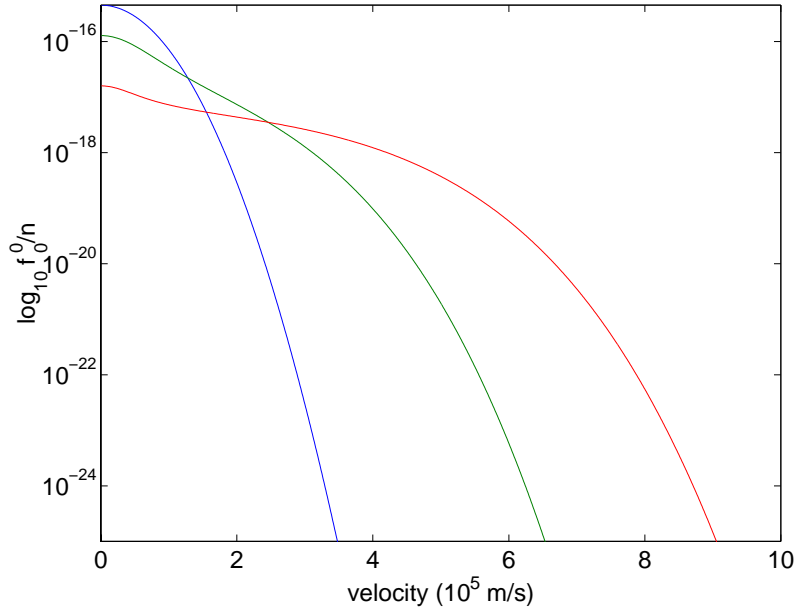
Figure 3-1: Distribution functions for various field amplitudes (0, 0.5, 1 V/m), 3.3 MHz, R-mode, neutral density of 10^{21} m^{-3}.

solve for $\frac{\partial f_0^0}{\partial v}/f_0^0$, and integrate as follows:

$$f_0^0(v,t) = A \exp\left\{-\int_0^v \frac{mv}{\kappa T + e^2(E^1)^2 p(t)\nu/3mR[(\omega \mp \omega_c)^2 + \nu^2]}\, dv\right\}. \qquad (3.3)$$

The constant A is chosen such that the distribution function integrates to n. The distribution function is essentially a Maxwellian $f_0^0 = A\exp[-\int_0^v (mv/\kappa T)\, dv]$ with a second term in the denominator which modifies the temperature, or width, of the distribution. The term in question goes to zero as $v \to 0$ due to the factor of ν. Thus the shape is roughly a Maxwellian, with a spreading in the mid-velocity range. The spreading becomes more pronounced with increasing electric field.

To visually confirm this interpretation, in Figure 3-1 we numerically evaluate Equation (3.3) for a range of heater electric fields E^1, a carrier frequency $\omega/2\pi = 3.3$ MHz, R-mode polarization, a neutral density $N = 10^{21}$ m^{-3} and neutral temperature $T = 180$ K. The assumption is slow modulation, so $p(t)$ is taken to be 1. The distribution functions can be used to verify the condition of weak anisotropy assumed in Chapter 2 [Eq. (2.8)]. *Gurevich* [1978] derives the second spherical harmonic for
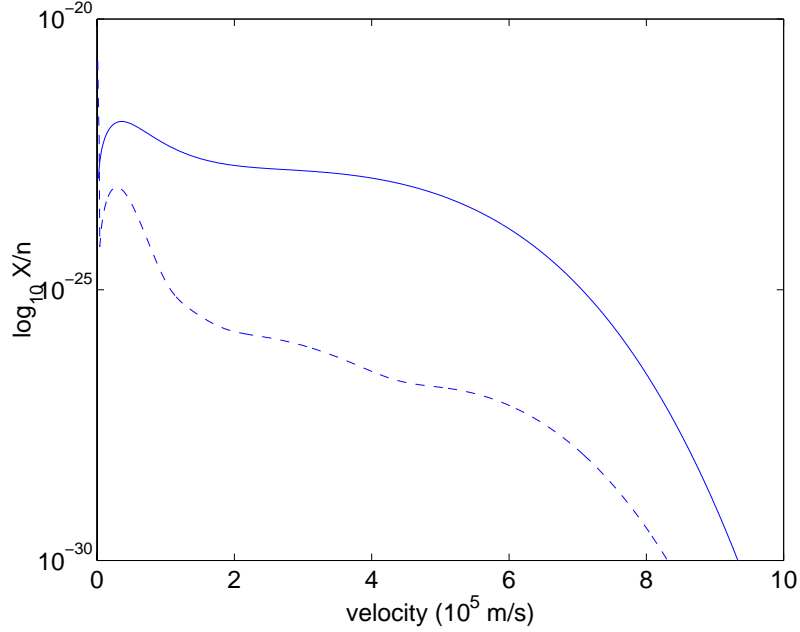
Figure 3-2: Left side (solid) and right side (dashed) of Equation (3.4).

the case of an unmagnetized plasma. It is found to negligible under the condition that

$$\left|\frac{\partial f_0}{\partial v}\right| \gg \left|\frac{4}{15v^3}\frac{\partial}{\partial v}\left\{\frac{eEv^4}{m(-i\omega+\nu)}\frac{\partial}{\partial v}\left[\frac{eE}{m(-i\omega+\nu)v}\frac{\partial f_0}{\partial v}\right]\right\}\right|. \tag{3.4}$$

We ignore the magnetic field here, and substitute the 1 V/m solution from Figure 3-1 into each side of Equation (3.4). Each side is plotted in Figure 3-2 (normalized to electron density, and indicated as X). We conclude that Equation (3.4) is well-satisfied at all velocities.

3.1.2 Approximate time constants

We now examine the range of time scales for which the slow modulation solution presented above is valid. The determination will be made on the basis of energy time constants. These are the initial rates of change of the energy in response to a step in $p(t)$, as shown in Figure 3-3. The heating and cooling rates from Equation (2.141)
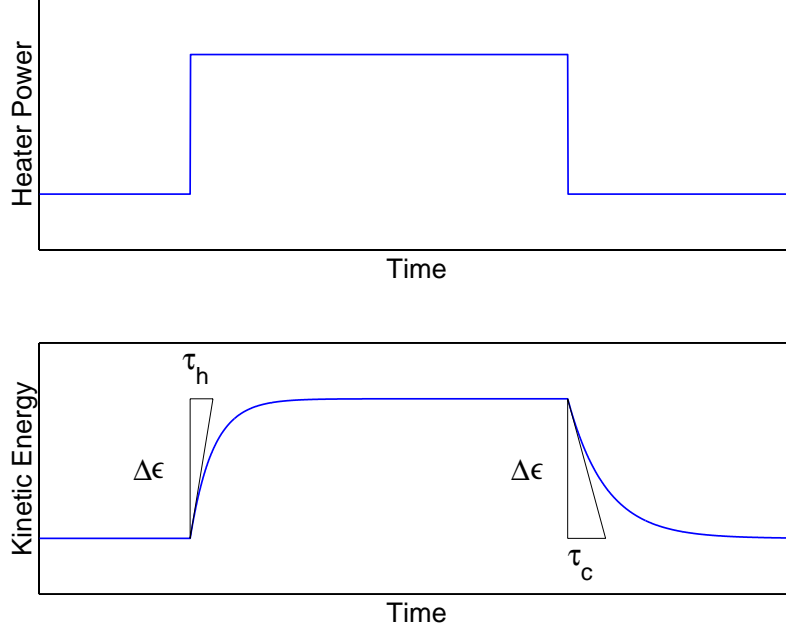
Figure 3-3: Definition of time constants.

are:

$$\frac{\partial \epsilon}{\partial t} = Q - L \tag{3.5}$$

$$Q = -\frac{2\pi e^2 (E^1)^2 p(t)}{3mn} \int \frac{v^3 \nu}{(\omega \mp \omega_c)^2 + \nu^2} \frac{\partial f_0^0}{\partial v} \, dv \tag{3.6}$$

$$L = \frac{2\pi m}{n} \int R v^4 \left(f_0^0 + \frac{\kappa T}{mv} \frac{\partial f_0^0}{\partial v} \right) dv. \tag{3.7}$$

In the case of heating, we assume the distribution is initially in equilibrium such that $L = 0$ [see Eq. (2.121)], and therefore the initial slope $\frac{\partial \epsilon}{\partial t}$ is simply Q, with $p(t) = 1$, and f_0^0 is Maxwell-distributed at the temperature of the molecules. Thus,

$$\tau_h = \frac{\Delta \epsilon}{Q}, \tag{3.8}$$

with $\Delta \epsilon$ given by the difference in energy of the final and initial distributions:

$$\Delta \epsilon = \frac{2\pi m}{n} \int [f_0^0 (t = \infty) - f_0^0 (t = 0)] v^4 \, dv. \tag{3.9}$$

60
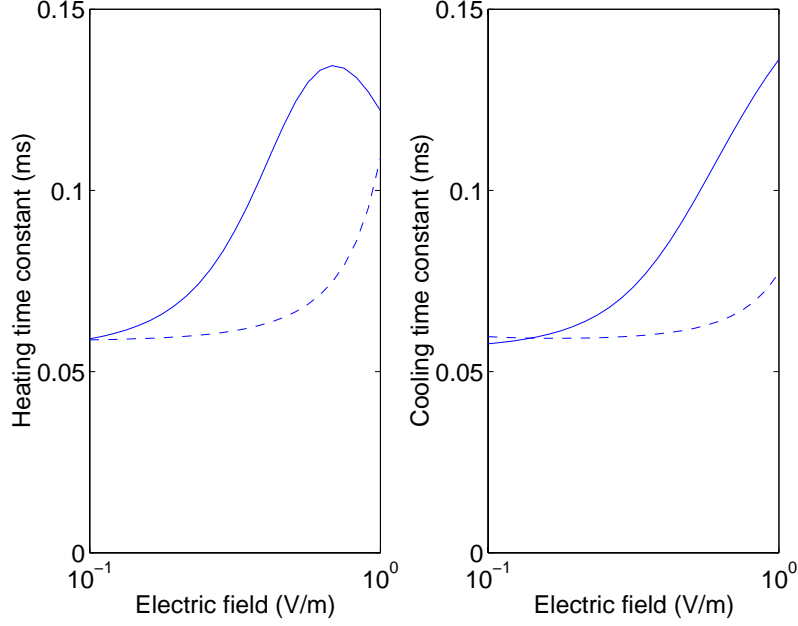
Figure 3-4: Heating, cooling time constants versus heater electric field, 3.3 MHz (solid), 10 MHz (dashed), R-mode polarization, neutral density of 10^{21} m^{-3}.

For cooling rates, we begin with a heated ionosphere in steady state and apply a falling step function $p(t)$. Immediately after the step, $p(t) = 0$ and the initial slope $\frac{\partial \epsilon}{\partial t}$ is just $-L$, with f_0^0 given by Equation (3.3). Thus,

$$\tau_c = \frac{\Delta \epsilon}{L}, \tag{3.10}$$

with $\Delta \epsilon$ again given by Equation (3.9). The above time constants are a function of heater electric field, neutral density, and carrier frequency. As an example, in Figure 3-4 we numerically evaluate τ_h and τ_c versus electric field for two different carrier frequencies (3.3 and 10 MHz), R-mode polarization, and a neutral density of 10^{21} m^{-3}. In all cases, the low field limit is 59 μs, and increases with field to about 130 μs at 1 V/m. The reason for the increase is that the rotational losses decrease with velocity, leading to a runaway effect. The runaway effect causes the temperature modification to increase rapidly with little additional heater power. In other words, the numerators of Equations (3.8) and (3.10) increase faster than the denominators. However, the elastic losses eventually catch up and limit the runaway effect. This

61

limit is seen in the levelling off of the heating time constant of 3.3 MHz heating near 1 V/m.

Thus in conclusion, the slow modulation limit is good up to 1–2 kHz for electric fields in the range 0.1–1 V/m.

3.1.3   Fast modulation limit

We now consider the limit where $p(t)$ varies much faster than τ (but still much slower than the HF carrier) such that f_0^0 can be characterized by small perturbations δf_0^0 around an average value $\overline{f_0^0}$, where $|\delta f_0^0| \ll |\overline{f_0^0}|$. In other words, the plasma does not have sufficient time to respond to temporal changes in heater power. In this limit, the product of $p(t)$ and f_0^0 in the Boltzmann equation (3.1) linearizes and the solution assumes a simple form.

To proceed, we expand the other time-varying quantity $p(t)$ in perturbation $\delta p(t)$ about a mean value $\overline{p(t)}$. However, the perturbation to $p(t)$ need not be small, that is $|\delta p(t)| \approx |\overline{p(t)}|$. Writing Equation (3.1) for the perturbed quantities gives us:

$$
\begin{aligned}
\frac{\partial}{\partial t}(\overline{f_0^0} + \delta f_0^0) &= \frac{e^2 (E^1)^2 [\overline{p(t)} + \delta p(t)]}{6m^2 v^2} \frac{\partial}{\partial v} \left[\frac{v^2 \nu}{(\omega \mp \omega_c)^2 + \nu^2} \frac{\partial}{\partial v}(\overline{f_0^0} + \delta f_0^0) \right] \\
&+ \frac{1}{2v^2} \frac{\partial}{\partial v} \left[Rv^3 \left(1 + \frac{\kappa T}{mv} \frac{\partial}{\partial v} \right) \right] (\overline{f_0^0} + \delta f_0^0).
\end{aligned} \tag{3.11}
$$

We break this expression into equations for the steady and perturbed quantities. For the steady quantities, we simply get the slow modulation limit [Eq. (3.3)] with f_0^0 replaced by $\overline{f_0^0}$, and $(E^1)^2 p(t)$ replaced by $(E^1)^2 \overline{p(t)}$. For the fluctuating quantities, we recognize the ordering $\delta p(t)\overline{f_0^0} \gg \overline{p(t)}\delta f_0^0 \approx \delta p(t)\delta f_0^0$, and arrive at

$$
\frac{\partial}{\partial t}\delta f_0^0 = \frac{e^2 (E^1)^2 \delta p(t)}{6m^2 v^2} \frac{\partial}{\partial v} \left[\frac{v^2 \nu}{(\omega \mp \omega_c)^2 + \nu^2} \frac{\partial \overline{f_0^0}}{\partial v} \right] + \frac{1}{2v^2} \frac{\partial}{\partial v} \left[Rv^3 \left(1 + \frac{\kappa T}{mv} \frac{\partial}{\partial v} \right) \right] \delta f_0^0.
$$
$$\tag{3.12}$$

Further simplification is possible by noticing that in the limit of rapidly fluctuating $\delta p(t)$, δf_0^0 goes to zero, and thus the loss term becomes negligible compared to the time derivative and heating terms. In this limit, we consider the response to 100%
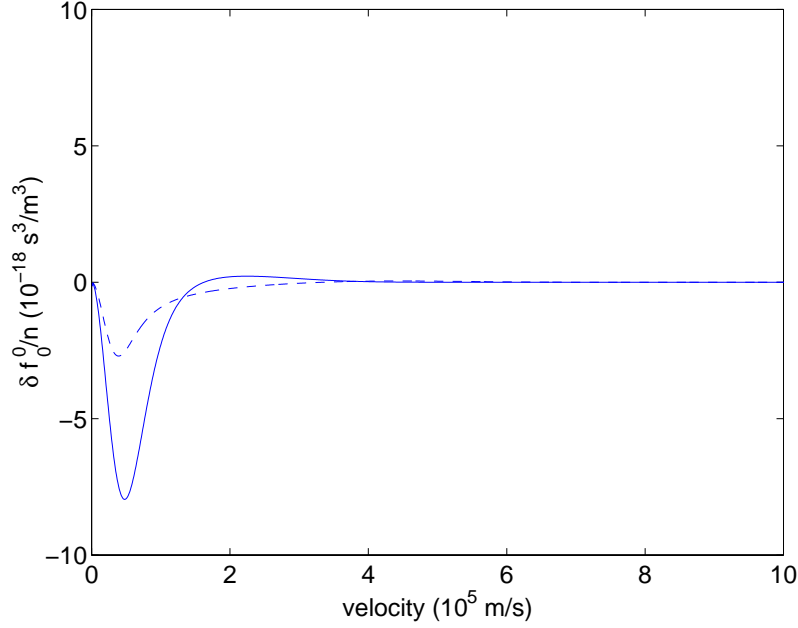
Figure 3-5: The fluctuating distributions for 0.5 V/m (solid) and 1 V/m (dashed), 3.3 MHz, R-mode, neutral density of 10^{21} m^{-3}.

sinusoidal power modulation $\delta p(t) = e^{-i\omega_m t}/2$, where ω_m is the modulation frequency. In this case:

$$\delta f_0^0 = \frac{ie^2(E^1)^2}{12\omega_m m^2 v^2} \frac{\partial}{\partial v} \left[\frac{v^2 \nu}{(\omega \mp \omega_c)^2 + \nu^2} \frac{\partial \overline{f_0^0}}{\partial v} \right]. \tag{3.13}$$

Thus the response to fast monochromatic power excitation is a distribution entirely determined by the average power $(E^1)^2 \overline{p(t)}$, which varies in time at a phase lag of $\pi/2$ with respect to the excitation. By extrapolation to a spectrum of frequencies, we conclude that the system integrates the fluctuating input power function in the fast modulation limit.

The fluctuating distribution functions $\delta f_0^0(v)$ are quite different than the steady-state distribution functions $\overline{f_0^0}$ found previously in that they are populated in areas where the first and second derivatives of the steady state solutions are nonzero. We plot them in Figure 3-5 for the same parameters as the steady-state case.

63

3.1.4 Variation with altitude

The solutions presented previously were for a single point in space with a neutral density of 10^{21} m^{-3}. However, the model needs to be expanded to allow for variations in neutral density with altitude. In particular, the neutral density falls off with increasing altitude, resulting in lower losses and more electron heating at higher altitudes. However, at the same time the electric field available for heating is less due to the inverse-square spreading and attenuation of the upgoing wave. These competing influences result in the heating maximum being located in the mid-range of altitudes.

The attenuation of the upward-going heater wave is nonlinear in the sense that it depends on the amplitude of the heater wave. Interpreted physically, the mechanism of plasma heating is self-reinforcing: the wave heats the plasma as it passes through, increasing the imaginary refractive index, leading to increased absorption and heating. In our case, the main consequence is that the heating becomes localized in a narrow range of altitudes, limiting the spatial extent of the antenna current.

In the ionosphere, R/L heater wave behaviour is determined by the parallel index of refraction:

$$n_\parallel^2 = 1 - \frac{\omega_p^2}{\omega(\omega + i\nu_e \mp \omega_c)}. \tag{3.14}$$

Here ν_e is an effective collision frequency [Eq. (2.53)] appearing in the momentum conservation fluid equation [Eq. (2.51)]. The refractive index arises from the high-frequency dielectric response of the plasma associated with $\mathbf{f}_1^1$ and thus we calculate an effective collision frequency as follows

$$\overline{\mathbf{v^1}}\nu_e = -\frac{4\pi}{3n} \int \mathbf{f}_1^1 \nu v^3 \, dv. \tag{3.15}$$

The replacement of $\overline{\mathbf{v^1}\nu}$ by $\overline{\mathbf{v^1}}\nu_e$ is possible when ν is isotropic in velocity space. In view of Equation (2.138) this approximation basically equates the effective and

averaged mobility tensor elements as follows:

$$\frac{\nu_e - i\omega}{\omega_c^2 + (\nu_e - i\omega)^2}\nu_e = -\frac{4\pi}{3n}\int \frac{\nu - i\omega}{\omega_c^2 + (\nu - i\omega)^2}\frac{\partial f_0^0}{\partial v}\nu v^3\, dv \qquad (3.16)$$

$$\frac{\omega_c}{\omega_c^2 + (\nu_e - i\omega)^2}\nu_e = -\frac{4\pi}{3n}\int \frac{\omega_c}{\omega_c^2 + (\nu - i\omega)^2}\frac{\partial f_0^0}{\partial v}\nu v^3\, dv \qquad (3.17)$$

$$\frac{1}{\nu_e - i\omega}\nu_e = -\frac{4\pi}{3n}\int \frac{1}{\nu - i\omega}\frac{\partial f_0^0}{\partial v}\nu v^3\, dv. \qquad (3.18)$$

With ω and ω_c known, ν_e can be found either by dividing Equations (3.16) and (3.17), or by dividing the real and imaginary parts of Equation (3.18). These two methods are thus:

$$\frac{\nu_e - i\omega}{\omega_c} = \left[\int \frac{(\nu - i\omega)\nu v^3}{\omega_c^2 + (\nu - i\omega)^2}\frac{\partial f_0^0}{\partial v}\, dv\right]\left[\int \frac{\omega_c \nu v^3}{\omega_c^2 + (\nu - i\omega)^2}\frac{\partial f_0^0}{\partial v}\, dv\right]^{-1} \qquad (3.19)$$

$$\frac{\nu_e}{\omega} = \left[\int \frac{\nu^2 v^3}{\nu^2 + \omega^2}\frac{\partial f_0^0}{\partial v}\, dv\right]\left[\int \frac{\nu \omega v^3}{\nu^2 + \omega^2}\frac{\partial f_0^0}{\partial v}\, dv\right]^{-1}. \qquad (3.20)$$

Numerical calculations will show that these two formulas produce almost identical results. The two methods (denoted as methods "one" and "two", respectively) are compared in Figure 3-6. To illustrate the character of ν_e, in Figure 3-7 we have plotted ν_e versus heater electric field for two different carrier frequencies (3.3 and 10 MHz), R-mode polarization, at a neutral density of 10^{21} m^{-3}, using the slow modulation approximation for f_0^0.

Returning to the discussion of refractive index, we note that in the lower ionosphere, HF transmissions are above the right-hand cutoff of the plasma, and therefore both modes are propagating. In fact, $\omega \gg \omega_p$, so the real part of the refractive index is close to one. Therefore we can expand the square on the left side of Equation (3.14) to give the imaginary part:

$$\Im(n_\parallel) = \frac{\omega_p^2 \nu_e}{2\omega[(\omega \mp \omega_c)^2 + \nu_e^2]}. \qquad (3.21)$$

Clearly, the imaginary refractive index will be large where ω_p^2 and ν_e are numerically of comparable size. With the index in hand, we can compute the heater power density
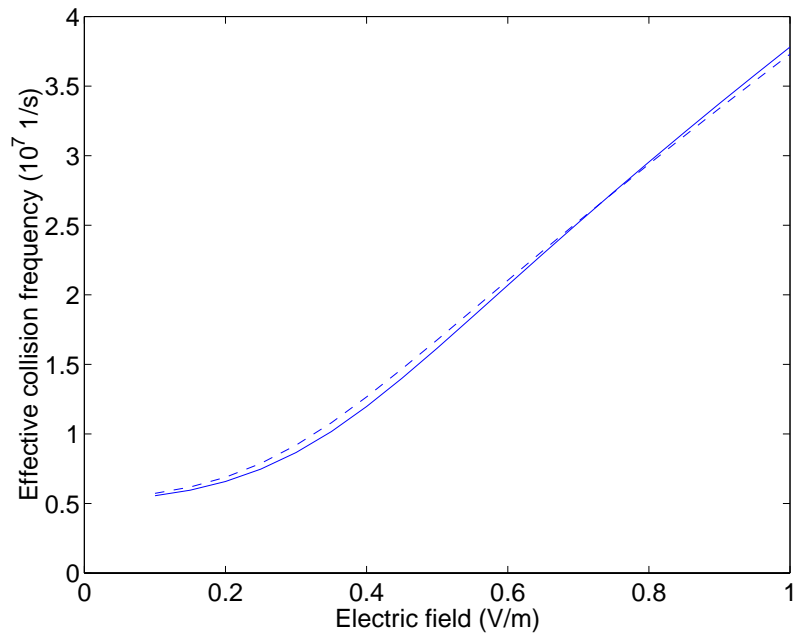
Figure 3-6: Comparison of methods to determine ν_e, method one (solid), method two (dashed).
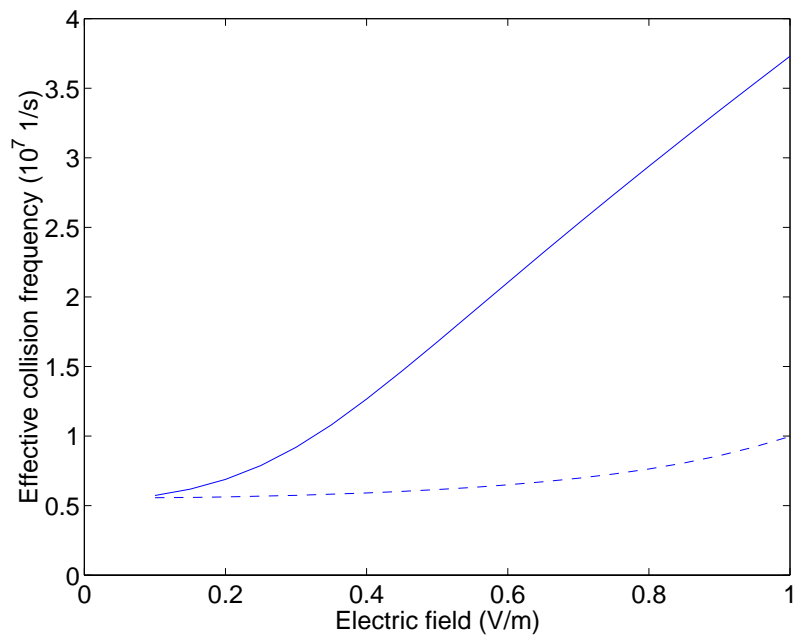

Figure 3-7: Effective collision frequency versus heater electric field, 3.3 MHz (solid), 10 MHz (dashed), R-mode, neutral density 10^{21} m^{-3}.

at the heating location:

$$\frac{(E^1)^2}{2\eta} = \frac{P_t G}{4\pi z^2} \exp\left[-\frac{2\omega}{c}\int_0^z \Im(n_\parallel)\, dz\right]. \tag{3.22}$$

Here, $P_t G$ is the transmitter power-gain product, often referred to as effective radiated power (ERP). To solve Equation (3.22) with a self-consistent electric field, the ionosphere is discretized into horizontal slabs. Equation (3.22) is used to calculate the field in the $(n+1)$th slab by considering attenuation in the first n slabs. We write this operation as follows:

$$\frac{(E^1_{n+1})^2}{2\eta} = \frac{P_t G}{4\pi z^2} \exp\left[-\frac{\Delta z}{c}\sum_{k=1}^n \frac{\omega_p^2 \nu_e}{(\omega \mp \omega_c)^2 + \nu_e^2}\right]. \tag{3.23}$$

where Δz is the slab thickness, and ω_p and ν_e vary with k. Solutions for E^1 as a function of altitude will depend on the electron density profile, the amount of power transmitted from the ground, and the chosen carrier frequency and polarization, as well as the particular limiting form of the distribution function [Equation (3.3) or (3.13)]. As an example, we will solve Equation (3.23) for the case of nighttime and daytime electron density profiles, assuming 20 MW ERP, a 3.3 MHz carrier in R-mode, and using the slow modulation approximation for f_0^0. The electron density profiles to be used represent normal levels of riometer absorption (see Chapter 1). As a second example, we choose a nighttime profile, 20 MW ERP, 3.3 and 10 MHz R-mode carrier, and again, slow modulation. These two examples are shown in Figure 3-8. Note that the daytime profile absorbs the heater electric field about 10 km lower than the nighttime profile due to the higher plasma density. In the second example, the 10 MHz field penetrates through to 100 km, whereas the 3.3 MHz field is absorbed at an altitude of about 70 km.

The heater electric field alone will not tell us the state of ionospheric modification. To determine modification, it is representative to plot temperature change $\Delta T_e(z) = 2\Delta\epsilon/3\kappa$, where $\Delta\epsilon$, the change in energy between on and off states of the heater, is given by Equation (3.9). Two plots have been prepared (Figures 3-9 and

67
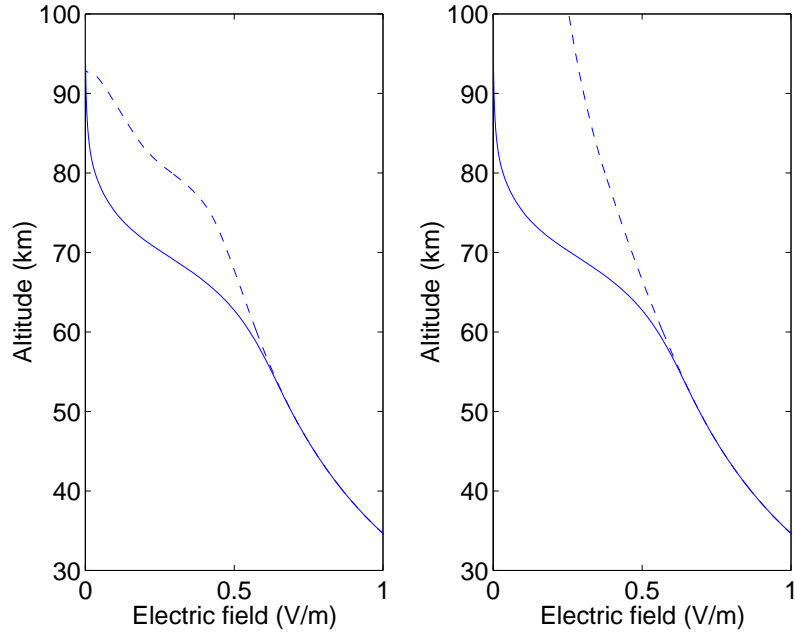
Figure 3-8: Heater electric field versus altitude. Left panel, daytime (solid), nighttime (dashed), 20 MW ERP, 3.3 MHz, R-mode. Right panel, daytime, 20 MW ERP, 3.3 MHz (solid), 10 MHz (dashed), R-Mode.

3-10) to show the variation with time of day and carrier frequency. Figure 3-9 shows the temperature for normal daytime and nighttime electron density profiles (moderate riometer absorption). The weaker nighttime profile results in less attenuation of the upgoing wave, so the heating is spread over a much broader altitude range, resulting in more modification. Thus the nighttime conditions are favored for heating experiments, particularly for modification of the plasma above 70 km altitude.

The variation with carrier frequency is given in Figure 3-10. The higher carrier frequencies suffer less attenuation, spreading the temperature excitation over a broader altitude range. However, the amount of heating near the peak is considerably less. Thus there is a tradeoff between penetration and heating efficiency which can be taken advantage of during the unfavorable daytime conditions. The details of the tradeoff will be made precise in the discussion of conductivity in the next section.
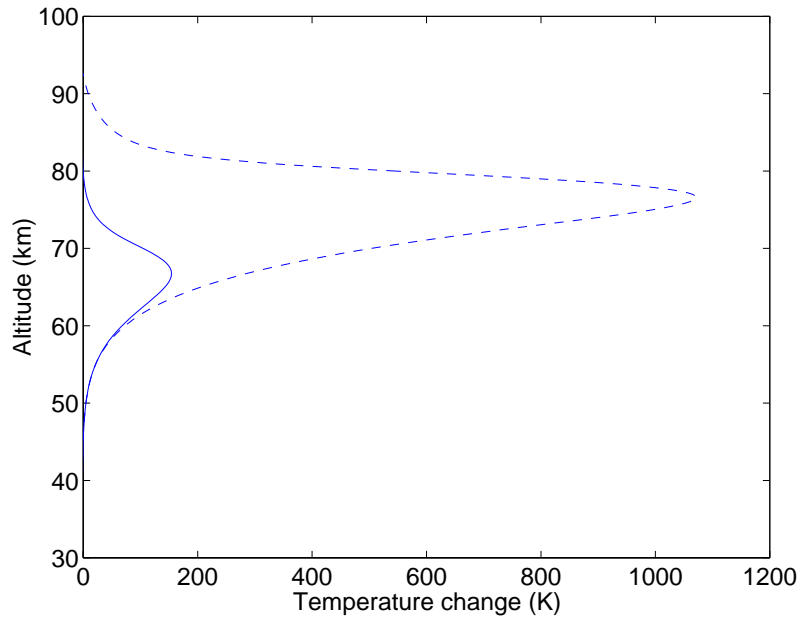
68

Figure 3-9: Temperature change profiles in daytime (solid), nighttime (dashed), 20 MW ERP, 3.3 MHz, R-mode, slow modulation limit.
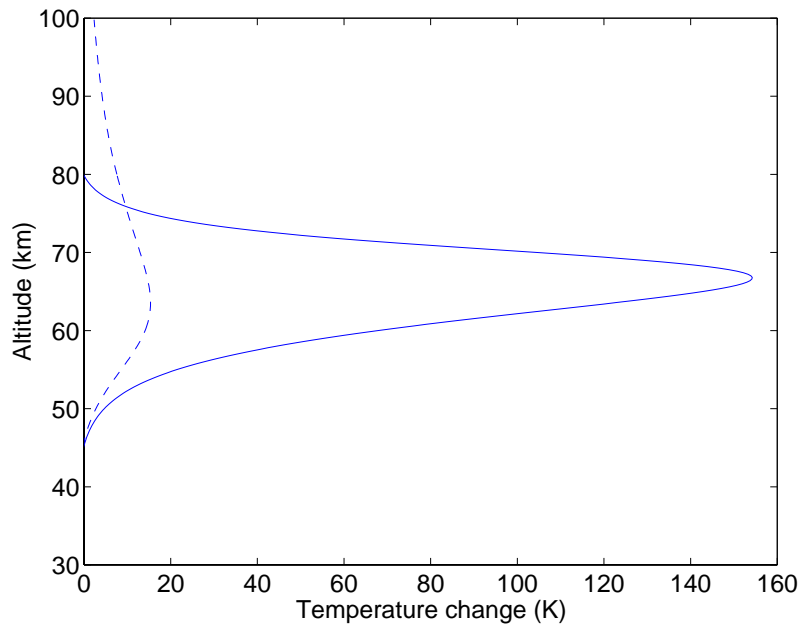

Figure 3-10: Temperature change profiles in daytime, 20 MW ERP, 3.3 MHz (solid), 10 MHz (dashed), R-mode, slow modulation limit.

3.2 Conductivity

The ionospheric conductivity follows directly from the distribution function via Equation (2.145):

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_P & -\sigma_H & 0 \\ \sigma_H & \sigma_P & 0 \\ 0 & 0 & \sigma_0 \end{bmatrix} \tag{3.24}$$

$$\sigma_P = -\frac{4\pi e^2}{3m} \int v^3 \frac{\nu}{\omega_c^2 + \nu^2} \frac{\partial f_0^0}{\partial v} \, dv \tag{3.25}$$

$$\sigma_H = -\frac{4\pi e^2}{3m} \int v^3 \frac{\omega_c}{\omega_c^2 + \nu^2} \frac{\partial f_0^0}{\partial v} \, dv \tag{3.26}$$

$$\sigma_0 = -\frac{4\pi e^2}{3m} \int v^3 \frac{1}{\nu} \frac{\partial f_0^0}{\partial v} \, dv. \tag{3.27}$$

The conductivity $\boldsymbol{\sigma}$ can be broken into two parts—a background conductivity $\overline{\boldsymbol{\sigma}}$ due to a Maxwellian f_0^0 with a temperature of the molecules, and a modulation (or conductivity change) $\Delta\boldsymbol{\sigma}$ due to the effect of the heater. The background electron conductivity is shown in Figure 3-11. The specific conductivity dominates over most of the altitude range. This supports the intuition that the field lines act as equipotentials.

As for the modulated conductivities, we note that the solar wind dynamo field $\mathbf{E}^0$ is applied perpendicular to the magnetic field (to the extent that the field lines can be regarded as equipotentials), and thus changes in σ_0 are unfelt. Our attention will be focused on $\Delta\sigma_P$ and $\Delta\sigma_H$, beginning with the slow modulation limit for f_0^0. Figures 3-12 and 3-13 show the conductivity changes for the cases presented in Figures 3-9 and 3-10. The left panel of each plot is the Pedersen conductivity change, and the right panel is the Hall conductivity change.

Figure 3-12 compares the daytime and nighttime cases. The trend is that the increased plasma density of the daytime case shifts the conductivity modulation down about 10 km, although the gross structure is preserved.

The Pedersen conductivity depends on $\nu/(\omega_c^2 + \nu^2)$, and thus the conductivity
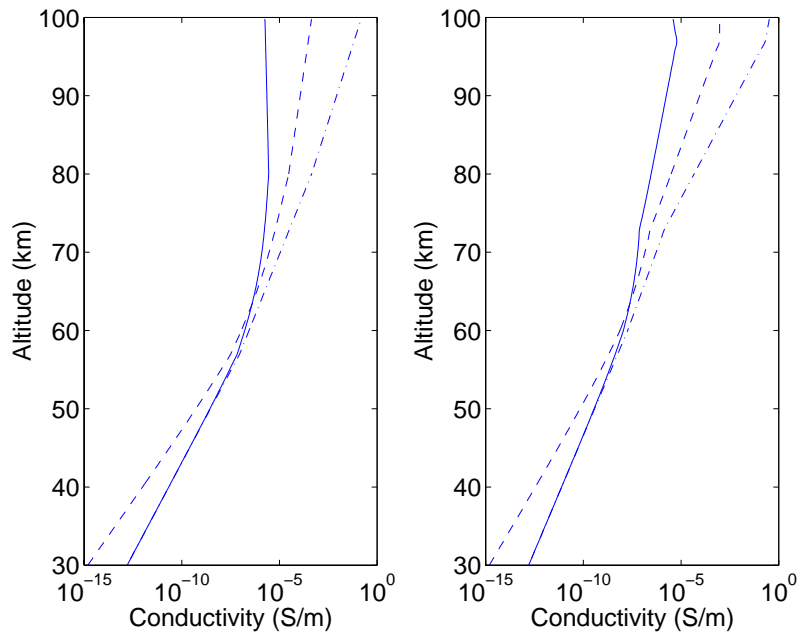
Figure 3-11: Background electron conductivity: Pedersen (solid), Hall (dashed), and specific (dot-dashed). Left panel is daytime, right panel is nighttime.
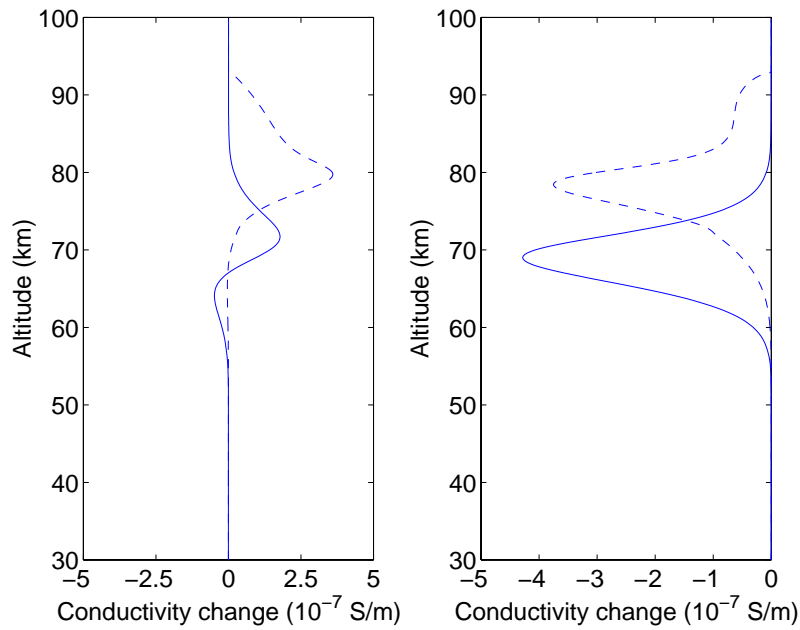
Figure 3-12: Conductivity change profiles in daytime (solid), nighttime (dashed), 20 MW ERP, 3.3 MHz, R-mode, slow modulation limit. Left panel is Pedersen and right panel is Hall.
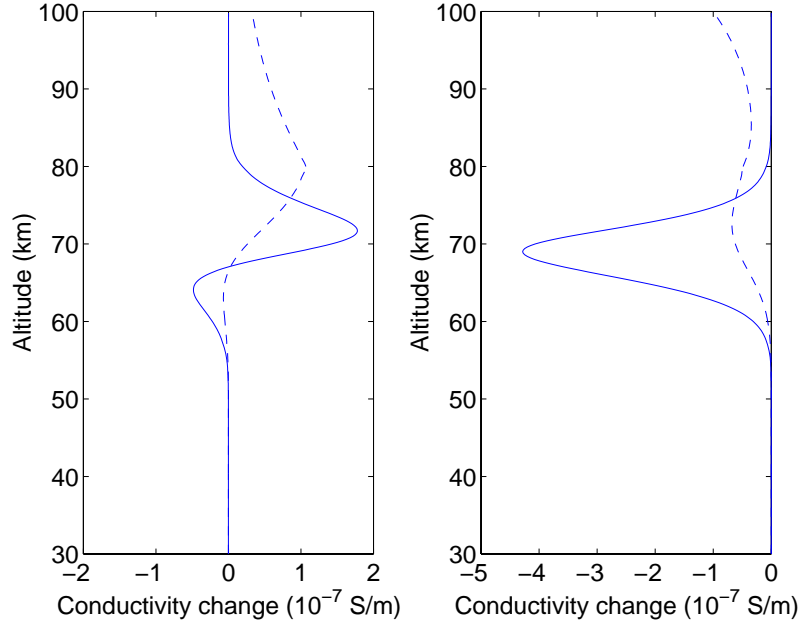
Figure 3-13: Conductivity change profiles in daytime, 20 MW ERP, 3.3 MHz (solid), 10 MHz (dashed), R-mode, slow modulation limit. Left panel is Pedersen and right panel is Hall.

change can be understood as follows. At high altitudes, the second term in the denominator can be ignored, and the conductivity depends directly on ν. In contrast, the first term in the denominator is disposable at low altitudes, resulting in a $1/\nu$ dependence. The sign reversal in the conductivity change occurs where the ω_c and ν denominator terms have comparable influence.

As for the Hall conductivity, the dominant feature is a negative conductivity change at all altitudes, since the dependence is $\omega_c/(\omega_c^2 + \nu^2)$. However, at high altitudes, the Hall conductivity modulation is rather inefficient compared to the Pedersen modulation due to the dominance of the ω_c^2 term in the denominator. Thus one can conclude generally that the Hall modulation is dominant at low altitudes and the Pedersen modulation is dominant at higher altitudes.

Figure 3-13 compares the conductivity changes for two different HF carrier frequencies: 3.3 and 10 MHz. The results suggest that the efficient Pedersen conductivity at higher altitudes can be accessed during the daytime conditions by increasing the carrier frequency. This becomes the dominant trend at higher heating powers
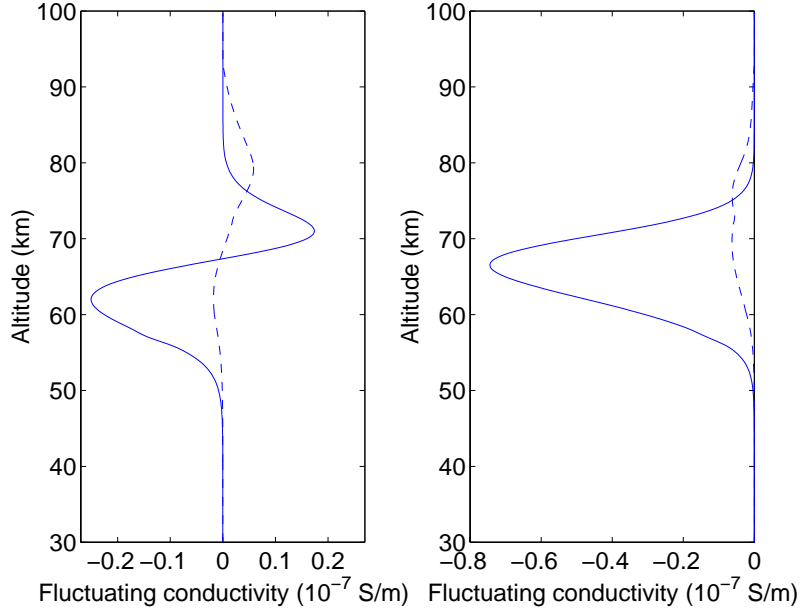
Figure 3-14: Fluctuating conductivity change profiles in daytime (solid), nighttime (dashed), 20 MW ERP, 3.3 MHz, R-mode, 10 kHz modulation frequency. Left panel is Pedersen and right panel is Hall.

(above 20 MW ERP), although the self-absorption effect eventually limits what can be achieved. Extrapolation to higher powers will not be discussed here as it is not directly verifiable by the author's experiments, and various projections have already been published [*Milikh et al.*, 1994, *Kuo et al.*, 2002].

In the case of fast modulation, the distribution function is no longer in steady state, and thus we think of the modulation as time-varying fluctuations around a mean modulation: $\Delta\sigma_{P,H} = \overline{\Delta\sigma_{P,H}} + \delta\sigma_{P,H}$. The $\overline{\Delta\sigma_{P,H}}$ term is associated with $\overline{f_0^0} - f_0^0[p(t) = 0]$, and is given by Figures 3-12 and 3-13. However, the $\delta\sigma_{P,H}$ term is associated with the fluctuating distribution function δf_0^0 [Eq. (3.13)] and needs to be evaluated by replacing f_0^0 with δf_0^0 in Equation (3.24). In Figures 3-14 and 3-15 we plot $\delta\sigma_{P,H}$ for the same cases as Figures 3-12 and 3-13.

Evidently, the conclusions from the slow modulation limit do not carry over to fast modulation. The difference between day and night is now more pronounced. While the profile still moves upward about 10 km, the conductivity modulation is much smaller in the nighttime case. The primary reason is that at higher altitudes, the
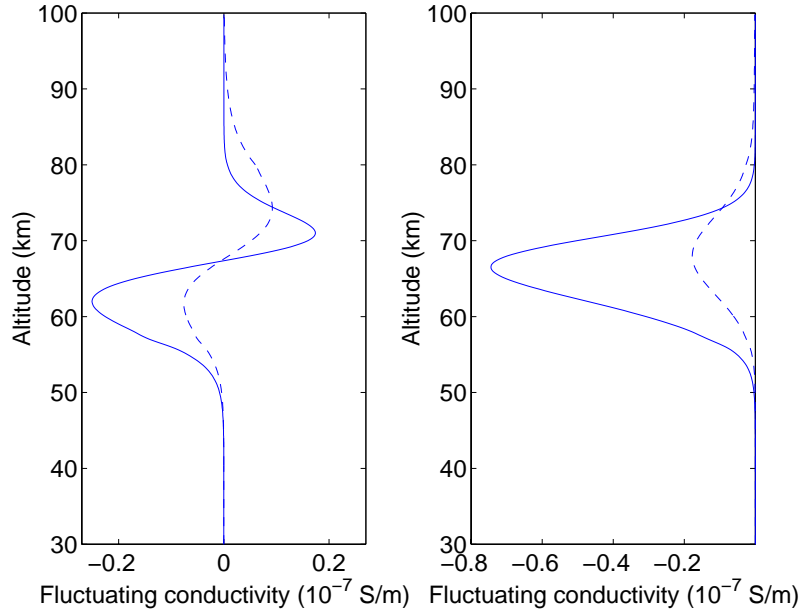
Figure 3-15: Fluctuating conductivity change profiles in daytime, 20 MW ERP, 3.3 MHz (solid), 10 MHz (dashed), R-mode, 10 kHz modulation frequency. Left panel is Pedersen and right panel is Hall.

neutral density decreases, causing the energy gain and loss rates to slow down, and thus lengthening the time constant τ. This time constant corresponds to the location of a pole in the response curve of conductivity change versus modulation frequency. As the pole moves to lower frequency, the conductivity response above the location of the pole gets pushed down.

Changing the carrier frequency is not helpful in this case for the same reason. As can be seen in Figure 3-15 the conductivity modulation, particularly the Pedersen conductivity, does not move to higher altitudes in any significant way. We conclude that operating at higher carrier frequencies is only helpful when the modulation frequency ω_m lies below $1/\tau$ at the altitudes being modulated. Viewed another way, the temporal response of the conductivity modulation process tends to localize effects at altitudes where $\omega_m \approx 1/\tau$.

3.3 Current

The current will be described in several levels of theoretical detail. The first level (Section 3.3.1) describes the theory of *Stubbe et al.* [1982], where the polarization of the modified plasma volume is ignored. The second level (Section 3.3.2) is the theory of *Stubbe and Kopka* [1977], which gives a two-dimensional analytic model for static polarization. The third level (Section 3.3.3) is the author's extension of the second level to three dimensions. The fourth level (Section 3.3.4) describes the author's numerical solution to Maxwell's equations.

3.3.1 Case of no polarization

The modification of the ionospheric conductivity perturbs the natural flow of electro-jet current in the ionosphere. The very simplest model of the perturbation considers the heated plasma to be homogeneous and infinite in extent. The current flowing in the plasma is given by

$$\mathbf{J} = \boldsymbol{\sigma}\mathbf{E}^0. \tag{3.28}$$

If the conductivity is varied by an amount $\Delta\boldsymbol{\sigma}$ around an average value of $\overline{\boldsymbol{\sigma}}$, then the current will similarly vary around an average value according to:

$$\overline{\mathbf{J}} + \Delta\mathbf{J} = (\overline{\boldsymbol{\sigma}} + \Delta\boldsymbol{\sigma})\mathbf{E}^0. \tag{3.29}$$

We are concerned with the varying current $\Delta\mathbf{J}$ as it is what will give rise to radiation.

We now choose a system of coordinates. Without loss of generality, the solar wind dynamo electric field $\mathbf{E}^0$ will point in the $\hat{\mathbf{x}}$ direction. To be consistent with the literature, $\hat{\mathbf{x}}$ will point east, $\hat{\mathbf{y}}$ will point north, and $\hat{\mathbf{z}}$ will point up (opposite the Earth's magnetic field). In this choice of coordinates, the conductivity tensor is given

by

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_P & \sigma_H & 0 \\ -\sigma_H & \sigma_P & 0 \\ 0 & 0 & \sigma_0 \end{bmatrix}. \tag{3.30}$$

$$\tag{3.31}$$

The difference from Equation (3.24) lies in the sign of the Hall conductivity. The current $\Delta \mathbf{J}$ is given by

$$\Delta \mathbf{J} = (\Delta \sigma_P \hat{\mathbf{x}} - \Delta \sigma_H \hat{\mathbf{y}}) E^0 \tag{3.32}$$

In this case, the perturbation current is oriented at an angle

$$\alpha = -\arctan\left(\frac{\Delta \sigma_H}{\Delta \sigma_P}\right) \tag{3.33}$$

counterclockwise (when viewed from above) with respect to $\mathbf{E}^0$.

At higher modulation frequencies, the conductivity change $\Delta \boldsymbol{\sigma}$ takes the form of small perturbations $\delta \boldsymbol{\sigma}$ around an average conductivity change $\overline{\Delta \boldsymbol{\sigma}}$. The steady and time-varying components are given by

$$\overline{\mathbf{J}} + \overline{\Delta \mathbf{J}} = (\overline{\boldsymbol{\sigma}} + \overline{\Delta \boldsymbol{\sigma}})\mathbf{E}^0 \tag{3.34}$$

$$\delta \mathbf{J} = \delta \boldsymbol{\sigma} \, \mathbf{E}^0. \tag{3.35}$$

If $\mathbf{E}^0$ is again assumed to be eastward, then

$$\delta \mathbf{J} = (\delta \sigma_P \hat{\mathbf{x}} - \delta \sigma_H \hat{\mathbf{y}}) E^0. \tag{3.36}$$

In this case, the perturbation current is oriented at an angle

$$\alpha = -\arctan\left(\frac{\delta \sigma_H}{\delta \sigma_P}\right) \tag{3.37}$$

counterclockwise (when viewed from above) with respect to $\mathbf{E}^0$.
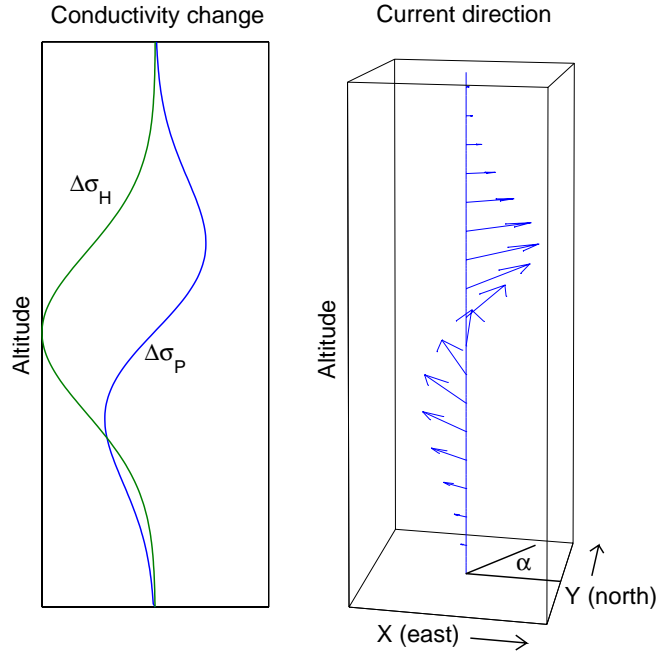
Figure 3-16: Variation of current direction with altitude, according to *Stubbe et al.* [1982].

These results can be simply extended to three dimensions by letting the conductivities become a function of space. This is the approach of *Stubbe et al.* [1982]. A qualitative view of the variation of current direction with altitude is given in Figure 3-16.

3.3.2 Two-dimensional polarization

The infinite, homogeneous model given above is only valid to the extent which the boundary of the heated volume can be ignored. In practice the heated volume is finite in extent, and develops polarization charge at its boundary in response to the need to keep current continuous across the conductivity discontinuity. This polarization charge is associated with a polarization electric field $\Delta\mathbf{E}$ which can drive additional current both inside and outside the heated volume. If $|\Delta\boldsymbol{\sigma}| \ll |\overline{\boldsymbol{\sigma}}|$, then the steady

and time-varying currents are given by

$$\overline{\mathbf{J}} = \overline{\boldsymbol{\sigma}}\mathbf{E}^0 \tag{3.38}$$

$$\Delta\mathbf{J} \approx \Delta\boldsymbol{\sigma}\mathbf{E}^0 + \overline{\boldsymbol{\sigma}}\Delta\mathbf{E}. \tag{3.39}$$

For rapidly fluctuating conductivity, the steady and time-varying components are given by

$$\overline{\mathbf{J}} + \overline{\Delta\mathbf{J}} \approx (\overline{\boldsymbol{\sigma}} + \overline{\Delta\boldsymbol{\sigma}})\mathbf{E}^0 + \overline{\boldsymbol{\sigma}}\,\overline{\Delta\mathbf{E}} \tag{3.40}$$

$$\delta\mathbf{J} \approx \delta\boldsymbol{\sigma}\,\mathbf{E}^0 + \overline{\boldsymbol{\sigma}}\,\delta\mathbf{E}. \tag{3.41}$$

Notice that while Equations (3.39) and (3.41) have essentially the same form, they involve perturbations around different equilibria [Equations (3.38) and (3.40), respectively], and therefore need to be treated distinctly. However, if both $|\overline{\Delta\boldsymbol{\sigma}}| \ll |\overline{\boldsymbol{\sigma}}|$ and $|\overline{\Delta\mathbf{E}}| \ll |\mathbf{E}^0|$, then $|\overline{\Delta\mathbf{J}}| \ll |\overline{\mathbf{J}}|$ and the equilibria can be considered the same. Thus $\delta\mathbf{J}$ can be determined from $\delta\boldsymbol{\sigma}$ using the same procedure that determines $\Delta\mathbf{J}$ from $\Delta\boldsymbol{\sigma}$.

We now proceed to determine $\Delta\mathbf{E}$, following the theory of *Stubbe and Kopka* [1977]. We consider the heated volume to be a vertical cylinder above the transmitter, aligned with the nearly-vertical magnetic field, as shown in Figure 3-17. The size of the cylinder is determined by the transmitter beam width. Because the solar dynamo field is applied horizontally, it is possible to integrate all quantities over height and reduce the three-dimensional cylinder model to a two-dimensional disc. In the height-integrated picture, the plasma has a height-integrated conductivity $\boldsymbol{\Sigma} = \int \boldsymbol{\sigma}\,dz$, which is associated with a height-integrated current density $\mathbf{K} = \int \mathbf{J}\,dz$. The conductivity inside the disc is $\boldsymbol{\Sigma}'$, and outside is $\boldsymbol{\Sigma}$. The geometry is shown in Figure 3-18. A vertical plasma column of conductivity different than the surroundings gives rise to a polarization electric field $\Delta\mathbf{E}$. We assume that the polarization is static and is associated with a potential U:

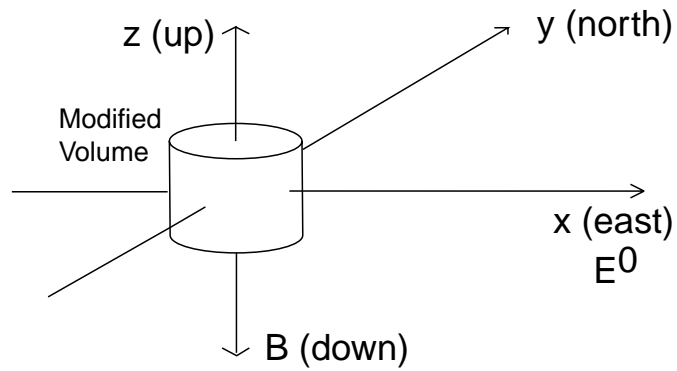$$\Delta\mathbf{E} = -\nabla U. \tag{3.42}$$
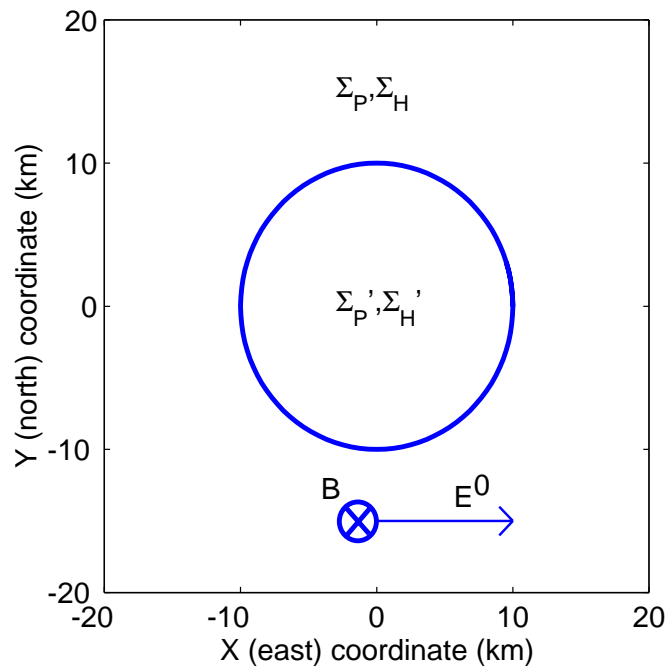
Figure 3-17: Cylinder of modified conductivity.

Figure 3-18: Coordinate system for a heated volume with a radius of 10 km.

In the polar coordinates of the disc, the total electric field is

$$E_r = E^0 \cos\phi - \frac{\partial U}{\partial r} \tag{3.43}$$

$$E_\phi = -E^0 \sin\phi - \frac{1}{r}\frac{\partial U}{\partial \phi}. \tag{3.44}$$

In view of the conductivity tensor, we write the polar components of the current as

$$K_r = \Sigma_P E_r + \Sigma_H E_\phi \tag{3.45}$$

$$K_\phi = \Sigma_P E_\phi - \Sigma_H E_r, \tag{3.46}$$

where Σ_P and Σ_H are the Pedersen and Hall components of the height-integrated conductivity tensor.

We assume $\nabla \cdot \mathbf{K} = 0$ (for static fields), which means that U obeys Laplace's equation inside and outside the disc:

$$\frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} + \frac{1}{r^2}\frac{\partial^2 U}{\partial \phi^2} = 0. \tag{3.47}$$

The most general solution is

$$U(r, \phi) = A_0 + \sum_{n=1}^{\infty}(A_n r^n + B_n r^{-n})\cos(n\phi - \alpha_n). \tag{3.48}$$

We match boundary conditions to solve for the coefficients. Outside the disc $(r \geq R)$, we need the potential to tend to zero at ∞, thus $A_n = 0$ for all n. Inside the disc, the solution must stay bounded, so $B_n = 0$ for all n. Furthermore, since the conductivities are rotationally invariant and the imposed field is antisymmetric upon a rotation by π, we expect that the current patterns will have similar π-rotation symmetry. Therefore only the A_1 and B_1 terms are nonzero. Since U must be continuous at the boundary of the region, $A_1 = B_1/R^2 \equiv D$. Using a primed/unprimed notation to denote quantities inside/outside the modified region, we have:

$$U'(r, \phi) = Dr\cos(\phi - \alpha) \tag{3.49}$$

$$U(r, \phi) = D\frac{R^2}{r}\cos(\phi - \alpha). \tag{3.50}$$

This can be thought of as the potential due to surface charge on the boundary of the modified disc. The polarization charge is rotated by an azimuthal angle α with respect to $\mathbf{E}^0$. The potential effectively steers the current such that it remains divergenceless.

The final boundary condition we will apply to the solution is that the current normal to the boundary K_r must be continuous. From Equation (3.45), we have

$$\Sigma'_P E'_r + \Sigma'_H E'_\phi = \Sigma_P E_r + \Sigma_H E_\phi. \tag{3.51}$$

Upon taking the appropriate derivatives of the potential functions and substituting them, we arrive at the following expressions for the parameters:

$$D = E^0 \left[\frac{\Delta\Sigma_P^2 + \Delta\Sigma_H^2}{(\Sigma'_P + \Sigma_P)^2 + \Delta\Sigma_H^2}\right]^{1/2} \tag{3.52}$$

$$\alpha = -\arctan\frac{(\Sigma'_P + \Sigma_P)\Delta\Sigma_H - \Delta\Sigma_P\Delta\Sigma_H}{(\Sigma'_P + \Sigma_P)\Delta\Sigma_P + \Delta\Sigma_H^2}, \tag{3.53}$$

where $\Delta\Sigma_{P,H} = \Sigma'_{P,H} - \Sigma_{P,H}$. In the limit of weak modification ($|\Delta\Sigma_{P,H}| \ll \Sigma_{P,H}$) these expressions simplify:

$$D = \frac{E^0}{2\Sigma_P}\left(\Delta\Sigma_P^2 + \Delta\Sigma_H^2\right)^{1/2} \tag{3.54}$$

$$\alpha = -\arctan\frac{\Delta\Sigma_H}{\Delta\Sigma_P}. \tag{3.55}$$

Therefore the currents inside and outside the cylinder are given by

$$K'_r = E^0(\Sigma'_P \cos\phi - \Sigma'_H \sin\phi) - D[\Sigma'_P \cos(\phi - \alpha) - \Sigma'_H \sin(\phi - \alpha)] \tag{3.56}$$

$$K'_\phi = -E^0(\Sigma'_P \sin\phi + \Sigma'_H \cos\phi) + D[\Sigma'_P \sin(\phi - \alpha) + \Sigma'_H \cos(\phi - \alpha)] \tag{3.57}$$

$$K_r = E^0(\Sigma_P \cos\phi - \Sigma_H \sin\phi) + D\frac{R^2}{r^2}[\Sigma_P \cos(\phi - \alpha) + \Sigma_H \sin(\phi - \alpha)] \tag{3.58}$$

$$K_\phi = -E^0(\Sigma_P \sin\phi + \Sigma_H \cos\phi) + D\frac{R^2}{r^2}[\Sigma_P \sin(\phi - \alpha) - \Sigma_H \cos(\phi - \alpha)]. \tag{3.59}$$

We separate these quantities into background currents $\overline{K_r} = E^0(\Sigma_P \cos\phi - \Sigma_H \sin\phi)$, $\overline{K_\phi} = -E^0(\Sigma_P \sin\phi + \Sigma_H \cos\phi)$, and modulation currents $\Delta K_r = K_r - \overline{K_r}$, $\Delta K_\phi = K_\phi - \overline{K_\phi}$ due to the presence of the modified plasma column. After expanding the $(\phi - \alpha)$ arguments, the modulation currents are

$$\Delta K_r' = D\Sigma \cos(\phi - \phi_0) \tag{3.60}$$

$$\Delta K_\phi' = -D\Sigma \sin(\phi - \phi_0) \tag{3.61}$$

$$\Delta K_r = D\Sigma \frac{R^2}{r^2} \cos(\phi - \phi_0) \tag{3.62}$$

$$\Delta K_\phi = D\Sigma \frac{R^2}{r^2} \sin(\phi - \phi_0), \tag{3.63}$$

with

$$\phi_0 = \arctan\frac{\Sigma_H}{\Sigma_P} + \alpha \tag{3.64}$$

$$\Sigma = \sqrt{\Sigma_P^2 + \Sigma_H^2}. \tag{3.65}$$

The above expressions simplify if we write them in a coordinate system (u, v) rotated from (x, y) by the angle ϕ_0. The cartesian vectors in (u, v) are:

$$\Delta K_u' = D\Sigma \tag{3.66}$$

$$\Delta K_v' = 0 \tag{3.67}$$

$$\Delta K_u = D\frac{R^2}{r^2}\Sigma \cos 2\phi \tag{3.68}$$

$$\Delta K_v = D\frac{R^2}{r^2}\Sigma \sin 2\phi. \tag{3.69}$$

Thus the modulation current is uniform in the disc and dipolar outside, as shown in Figure 3-19.

3.3.3   Extension to three dimensions

The previous formulation ignored vertical variations, allowing the use of height-integrated conductivities. However, this model ionosphere confines the resulting cur-
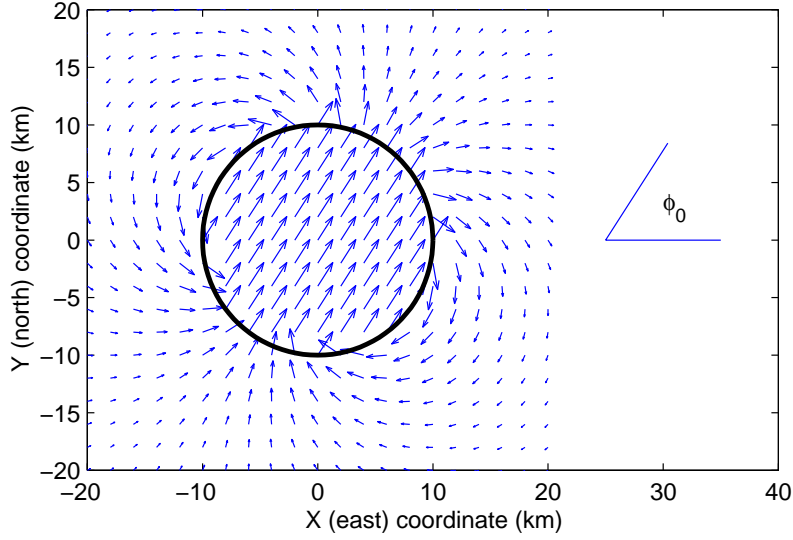
Figure 3-19: Modulation current pattern.

rent structure to a horizontal plane. In the real ionosphere, the specific conductivity is large, so currents can flow easily in the vertical direction. Furthermore, the Pedersen and Hall conductivities increase exponentially with altitude. Thus it is reasonable to hypothesize that the closure currents may in fact flow up and over the heated volume, rather than around the sides as in the previous analysis.

The model in the previous section needs to be reinterpreted in terms of the possibility of vertical current structure. As a first step, let us rewrite Equation (3.51), recognizing two distinct regions in altitude. The first region, at the level of the heated volume, has a conductivity discontinuity in the horizontal plane, as demarcated by the boundary between the primed (heated) and unprimed (unheated) areas. The second region consists of the unmodified plasma above the heated volume, where the conductivity is uniform in the horizontal plane. We will denote this region by using a double-primed notation. Assuming that current can flow easily along the field lines, Equation (3.51) can be written as:

$$(\Sigma'_P + \Sigma''_P)E'_r + (\Sigma'_H + \Sigma''_H)E'_\phi = (\Sigma_P + \Sigma''_P)E_r + (\Sigma_H + \Sigma''_H)E_\phi. \qquad (3.70)$$

The origin of this expression is shown in Figure 3-20.

$\Sigma''E'$ $\Sigma''E$

Current flows easily in vertical direction

$\Sigma'E'$ ΣE

Modified Volume
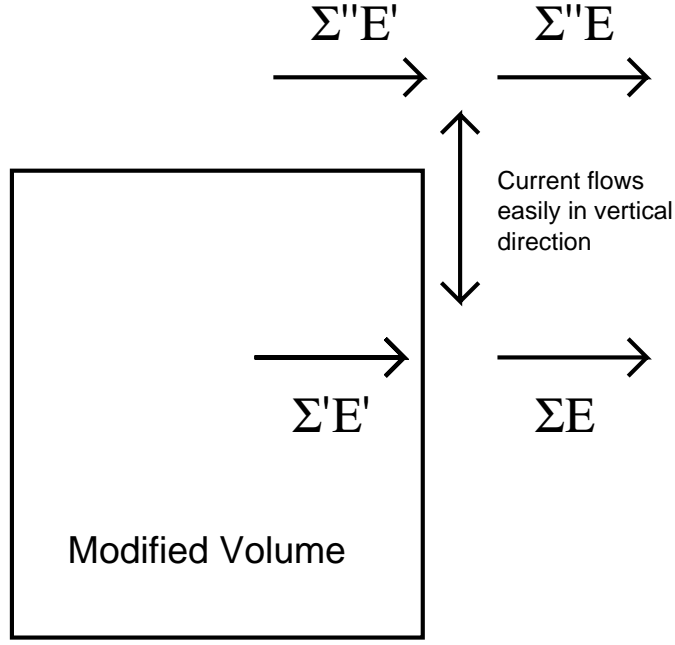
Figure 3-20: Currents forming the continuity relation [Eq. (3.70)].

After substituting Equations (3.43), (3.44), (3.49), and (3.50), we find that the amplitude and angle of the potential is given by:

$$D \;=\; E^0 \left[\frac{\Delta\Sigma_P^2 + \Delta\Sigma_H^2}{(\Sigma_P' + \Sigma_P + 2\Sigma_P'')^2 + \Delta\Sigma_H^2} \right]^{1/2} \tag{3.71}$$

$$\alpha \;=\; -\arctan \frac{(\Sigma_P' + \Sigma_P + 2\Sigma_P'')\Delta\Sigma_H - \Delta\Sigma_P \Delta\Sigma_H}{(\Sigma_P' + \Sigma_P + 2\Sigma_P'')\Delta\Sigma_P + \Delta\Sigma_H^2}, \tag{3.72}$$

where $\Delta\Sigma_{P,H} = \Sigma'_{P,H} - \Sigma_{P,H}$. In the limit of weak modification ($|\Delta\Sigma_{P,H}| \ll \Sigma_{P,H}$), and large conductivity gradient ($\Sigma_P'' \gg \Sigma_P$, see Figure 3-11), the previous expressions simplify:

$$D \;=\; \frac{E^0}{2\Sigma_P''} \left(\Delta\Sigma_P^2 + \Delta\Sigma_H^2 \right)^{1/2} \tag{3.73}$$

$$\alpha \;=\; -\arctan \frac{\Delta\Sigma_H}{\Delta\Sigma_P}. \tag{3.74}$$

Carrying out the same process as before to find the modulation currents, and contin-

uing to assume $\Sigma_P'' \gg \Sigma_P$, we find that at the altitude of the heated volume:

$$\Delta K_r' = 2D\Sigma_P'' \cos(\phi - \alpha) \tag{3.75}$$

$$\Delta K_\phi' = -2D\Sigma_P'' \sin(\phi - \alpha) \tag{3.76}$$

$$\Delta K_r = D\Sigma \frac{R^2}{r^2} \cos(\phi - \phi_0) \tag{3.77}$$

$$\Delta K_\phi = D\Sigma \frac{R^2}{r^2} \sin(\phi - \phi_0). \tag{3.78}$$

where Σ and ϕ_0 are as defined by Equations (3.64) and (3.65). Since $\phi_0 \neq \alpha$ and $\Sigma \neq 2\Sigma_P''$, the radial currents are no longer continuous at the boundary. The discontinuity in the radial current must be equal to the vertical-flowing current in order to ensure $\nabla \cdot \mathbf{J} = 0$. Furthermore, since $2\Sigma_P'' \gg \Sigma$, the currents inside the heated volume dominate the currents outside, and thus almost all the interior current is sent up the field lines upon reaching the boundary of the heated volume. For most purposes we can then ignore the exterior currents, and rewrite the system in cartesian coordinates (u, v) rotated from (x, y) by angle α:

$$\Delta K_u' = 2D\Sigma_P'' \tag{3.79}$$

$$\Delta K_v' = 0 \tag{3.80}$$

$$\Delta K_u = 0 \tag{3.81}$$

$$\Delta K_v = 0. \tag{3.82}$$

In the region above the heated volume, the modulation current solutions can be written as

$$\Delta K_r' = D\Sigma'' \cos(\phi - \phi_0) - 2D\Sigma_P'' \cos(\phi - \alpha) \tag{3.83}$$

$$\Delta K_\phi' = -D\Sigma'' \sin(\phi - \phi_0) + 2D\Sigma_P'' \sin(\phi - \alpha) \tag{3.84}$$

$$\Delta K_r = D\Sigma'' \frac{R^2}{r^2} \cos(\phi - \phi_0) \tag{3.85}$$

$$\Delta K_\phi = D\Sigma'' \frac{R^2}{r^2} \sin(\phi - \phi_0). \tag{3.86}$$
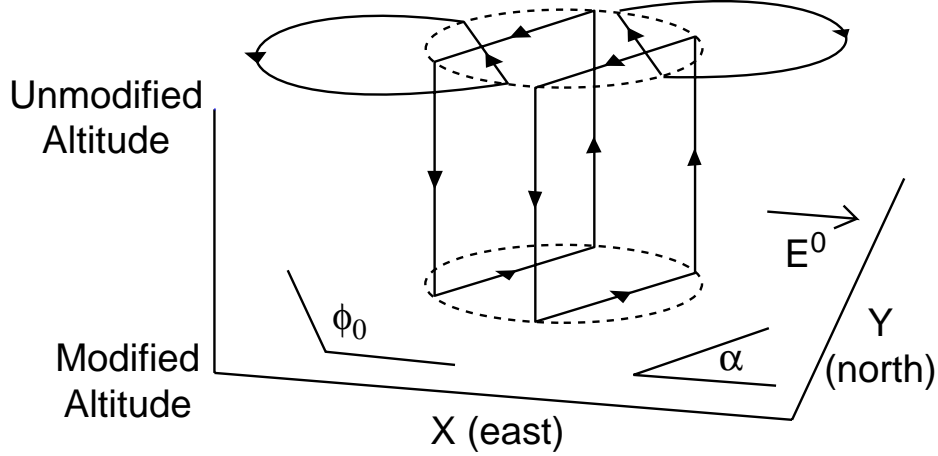
85

Figure 3-21: Combined horizontal and vertical modulation current structure.

with ϕ_0 as before, and $\Sigma'' = \sqrt{\Sigma_P''^2 + \Sigma_H''^2}$. The terms containing ϕ_0 are essentially the modulation currents of Equations (3.60)–(3.63), which describe a divergenceless horizontal current structure angled at ϕ_0 with respect to $\mathbf{E}^0$. The terms containing α can be written in cartesian coordinates (u, v) rotated from (x, y) by angle α:

$$\Delta K_u' = -2D\Sigma_P'' \qquad (3.87)$$

$$\Delta K_v' = 0 \qquad (3.88)$$

$$\Delta K_u = 0 \qquad (3.89)$$

$$\Delta K_v = 0. \qquad (3.90)$$

These are just the negatives of Equations (3.79)–(3.82). Continuity of these two systems require that $2D\Sigma_P''$ be the current flowing vertically along the field lines. The combined systems of current (3.79)–(3.82) and (3.83)–(3.86) are shown pictorially in Figure 3-21.

Although Figure 3-21 suggests a vertical loop of current, determining the actual size of the loop is nontrivial, since the Pedersen conductivity increases exponentially with altitude (see Figure 3-11). The higher altitudes tend to "short out" the current flowing at lower altitudes. One solution is to invoke magnetic induction, which can effectively truncate the vertical extent of the potential U driving the return current
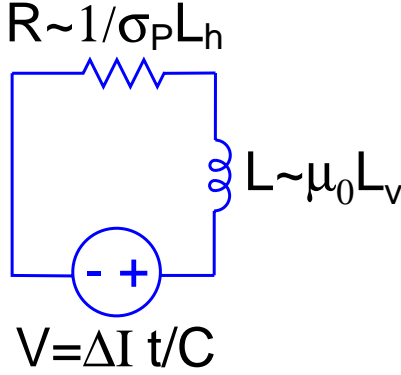
R~1/σ_PL_h

L~μ_0L_v

- +

V=ΔI t/C

Figure 3-22: Circuit model of vertical loop.

in the region above the heated volume. The approach will be described qualitatively here, with quantitative details given in Appendix A.

The basic intuition is that the vertical loop grows until the impedance due to the loop inductance is equal to the crossfield conduction impedance. A simple circuit model is provided in Figure 3-22. The voltage source driving the loop is simply the charge due to the conductivity discontinuity $Q = \int \Delta I \, dt$ divided by the capacitance of the heated volume. The loop inductance L is the permeability μ_0 times a vertical scale size L_v. The resistance of the return path R is approximately the reciprocal of the product of the Pedersen conductivity σ_P and a horizontal scale size L_h. These two impedances will be equal when the modulation time scale $1/\omega_m$ is equal to a magnetic diffusion time given by:

$$1/\omega_m = \tau_{mv} = \mu_0 \sigma_P L_h L_v, \tag{3.91}$$

which we will denote as the "vertical" magnetic time constant. If σ_P increases exponentially with altitude, then the vertical loop size will depend on the logarithm of ω_m.

An expanded lumped circuit model is given by Figure 3-23. Here we have added loops in the horizontal direction at both the heated altitude and above the heated altitude, with associated "horizontal" magnetic time constants τ'_{mh} and τ_{mh}, respectively. Whether they form at all will depend on the size of $1/\omega_m$ compared with

87
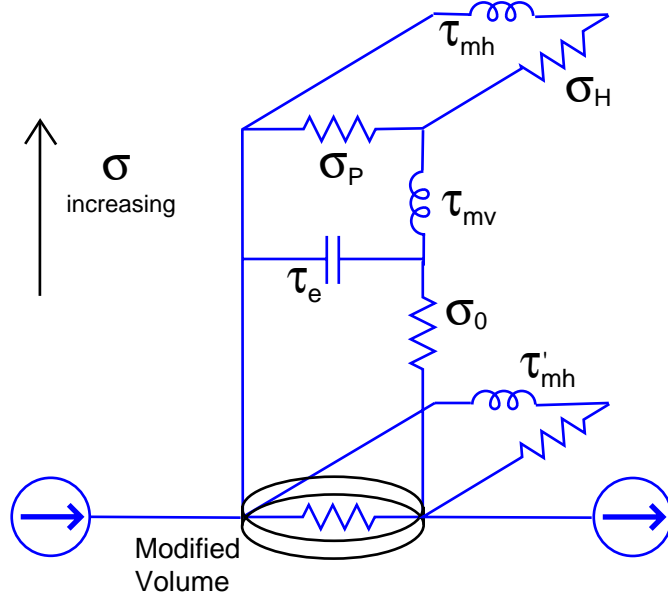
Figure 3-23: Circuit model of modified ionosphere.

τ'_{mh} and τ_{mh}. For completeness, an electric time constant τ_e is also considered in the vertical direction, to allow for a finite amount of time for charge to relax in the vertical direction. The electrojet current is modelled by current sources, and some of the dominant conductivities have been suggested in the circuit diagram. This circuit interpretation lays the framework for understanding the results of the next section.

3.3.4   Three-dimensional numerical model

The ideas presented in the previous section can be made concrete by a numerical solution to Maxwell's equations:

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \tag{3.92}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{3.93}$$

We can write the electric field as the sum of the solar dynamo field $\mathbf{E}^0$ and the polarization field $\Delta \mathbf{E}$ arising from the conductivity perturbation:

$$\mathbf{E} = \mathbf{E}^0 + \Delta \mathbf{E}. \tag{3.94}$$

The magnetic field consists of the steady magnetic field $\overline{\mathbf{H}}$ due to the steady current $\overline{\mathbf{J}}$ and a perturbation field $\Delta\mathbf{H}$ associated with the current perturbation $\Delta\mathbf{J}$:

$$\mathbf{H} = \overline{\mathbf{H}} + \Delta\mathbf{H}. \tag{3.95}$$

The earth's magnetic field does not enter the formulation because it is divergenceless, irrotational, time-constant, and is already accounted for in $\boldsymbol{\sigma}$. The current $\mathbf{J}$ is the combination of steady and perturbed currents given by Equations (3.38) and (3.39), such that

$$\mathbf{J} = \overline{\boldsymbol{\sigma}}\mathbf{E}^0 + (\Delta\boldsymbol{\sigma}\mathbf{E}^0 + \overline{\boldsymbol{\sigma}}\Delta\mathbf{E}). \tag{3.96}$$

Since Maxwell's equations are linear, the steady quantities $\mathbf{E}^0$, $\overline{\mathbf{H}}$, and $\overline{\boldsymbol{\sigma}}\mathbf{E}^0$ form a solution to Maxwell's equations which is steady in time. That solution can be removed from the equations, and we can re-solve the equations for the perturbed quantities

$$\nabla \times \Delta\mathbf{E} = -\mu_0\frac{\partial}{\partial t}\Delta\mathbf{H} \tag{3.97}$$

$$\nabla \times \Delta\mathbf{H} = \Delta\boldsymbol{\sigma}\,\mathbf{E}^0 + \overline{\boldsymbol{\sigma}}\Delta\mathbf{E} + \epsilon_0\frac{\partial}{\partial t}\Delta\mathbf{E}. \tag{3.98}$$

This system can be solved in time domain as follows. The system can be written as six scalar equations

$$-\mu_0\frac{\partial \Delta H_x}{\partial t} = \frac{\partial \Delta E_z}{\partial y} - \frac{\partial \Delta E_y}{\partial z} \tag{3.99}$$

$$-\mu_0\frac{\partial \Delta H_y}{\partial t} = \frac{\partial \Delta E_x}{\partial z} - \frac{\partial \Delta E_z}{\partial x} \tag{3.100}$$

$$-\mu_0\frac{\partial \Delta H_z}{\partial t} = \frac{\partial \Delta E_y}{\partial x} - \frac{\partial \Delta E_x}{\partial y} \tag{3.101}$$

$$\epsilon_0\frac{\partial \Delta E_x}{\partial t} = \frac{\partial \Delta H_z}{\partial y} - \frac{\partial \Delta H_y}{\partial z} - \Delta\sigma_P E^0 - \sigma_P\Delta E_x - \sigma_H\Delta E_y \tag{3.102}$$

$$\epsilon_0\frac{\partial \Delta E_y}{\partial t} = \frac{\partial \Delta H_x}{\partial z} - \frac{\partial \Delta H_z}{\partial x} + \Delta\sigma_H E^0 + \sigma_H\Delta E_x - \sigma_P\Delta E_y \tag{3.103}$$

$$\epsilon_0\frac{\partial \Delta E_z}{\partial t} = \frac{\partial \Delta H_y}{\partial x} - \frac{\partial \Delta H_x}{\partial y} - \sigma_0\Delta E_z. \tag{3.104}$$

The fields are given an initial condition $\Delta \mathbf{E}(t = 0)$ and $\Delta \mathbf{H}(t = 0)$ and the fields for $t > 0$ are determined by finite-differencing the space and time derivatives and iterating the equations. If the conductivity is zero, then the method is stable when the time discretization Δt and space discretizations $(\Delta x, \Delta y, \Delta z)$ satisfy the relation [*Yee*, 1966]:

$$\frac{\Delta t}{\sqrt{\epsilon_0 \mu_0}} < \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}. \tag{3.105}$$

When the conductivity is inhomogeneous (and anisotropic, as in this case) a rigorous stability criterion is difficult to establish.

The initial conditions are taken to be zero everywhere, with the drive being applied via the conductivity change $\Delta \sigma_{P,H}$. To simulate the experimental conditions, one would ideally apply a temporally sinusoidal conductivity drive at a given modulation frequency, and let the system evolve until it reaches the sinusoidal steady state. The simulation would then be repeated for a whole range of modulation frequencies covering the band of interest (say, 1–30 kHz). The frequency response at every point in space would then be known. In practice this procedure requires much computer time. For the purpose of understanding the experiments, sufficient knowledge of the ionospheric behaviour can be obtained by determining the high and low frequency limits of the frequency response. This information can be gleaned from the step response of the system. For example, if $g(t)$ is the system step response, $h(t)$ is the impulse response, and $\tilde{g}(\omega)$ and $\tilde{h}(\omega)$ are their Laplace transforms, then the initial and final value theorems state that

$$\lim_{\omega \to \infty} \tilde{h}(\omega) = \lim_{t \to 0} g(t) \tag{3.106}$$

$$\lim_{\omega \to 0} \tilde{h}(\omega) = \lim_{t \to \infty} g(t). \tag{3.107}$$

Thus the fast and slow scale behaviour of the step response will give us the high and low frequency behaviour of the system function.

Accordingly, we apply a conductivity profile $\delta \sigma_{P,H}$ at time zero, and find the transient response of the fields $\Delta \mathbf{E}$ and $\Delta \mathbf{H}$ with their initial state being zero. We choose the profile $\delta \sigma_{P,H}$ rather than $\Delta \sigma_{P,H}$ because we want to capture the ionospheric

response to a conductivity change at time scales faster than the thermal response time τ of the ionosphere.

The simulations were run in two and three dimensions. The two dimensional code attempts to verify the two-dimensional predictions of the direction of the perturbed current inside the heated volume for the case of no polarization electric fields [Eqs. (3.33) and (3.37)] and fully-developed polarization fields [Eq. (3.64)]. Two-dimensional simulations were run for daytime density/conductivity profiles in a 50×50 km plane, discretized into 1-km squares. The temporal discretization was taken as 2 μs. The boundary conditions for the simulation volume are periodic, which is adequate so long as the solution does not impinge on the boundaries of the box. The code was run for the conductivity parameters

$$\sigma_P = 1 \times 10^{-6} \text{ S/m} \tag{3.108}$$

$$\sigma_H = 2 \times 10^{-6} \text{ S/m} \tag{3.109}$$

$$\delta\sigma_P = 1 \times 10^{-8} \text{ S/m} \tag{3.110}$$

$$\delta\sigma_H = -2 \times 10^{-8} \text{ S/m.} \tag{3.111}$$

The background conductivities are the levels at 68 km during the day, and the modified conductivities are arbitrary examples. The heated volume was assumed to be of a Gaussian-distributed shape with a scale size of 20 km at the center of the simulation box.

The output of the simulation is the transient response of the ionosphere. The character of the current system is very different at short and long time scales. The short time scale behaviour (at 2 μs) is given in Figure 3-24. At short time scales, the current direction resolves to the angle prediction in the absence of polarization electric field [Eq. (3.37)]:

$$\alpha = -\arctan\left(\frac{\delta\Sigma_H}{\delta\Sigma_P}\right). \tag{3.112}$$

For our choice of parameters, α is 63.4 degrees. This angle is reflected in the direction of the current inside the heated disc with respect to the $\hat{\mathbf{x}}$ direction. There is no
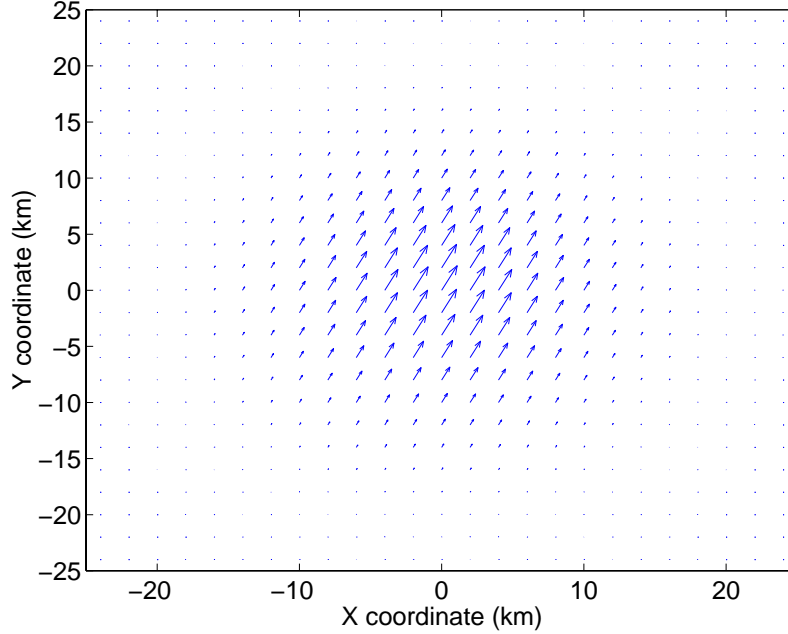
Figure 3-24: Current pattern at 2 μs.

current in the region outside the disc, thus there exists no significant polarization electric field. All of the perturbed current is due to the conductivity modification alone:

$$\delta \mathbf{J} \approx \delta \boldsymbol{\sigma} \, \mathbf{E}^0. \tag{3.113}$$

We interpret this result to indicate that circulating polarization current outside the heated volume takes a finite amount of time to develop. The reason lies in the finite amount of time required for the magnetic field to diffuse in the horizontal direction into the region enclosed by the circulating currents. This is τ_{mh}, the horizontal magnetic time constant.

We would expect at time scales much longer than τ_{mh} that horizontal loops could form as predicted by the analytic theory of Section 3.3.2. In Figure 3-25 is plotted the perturbed current pattern at 2 ms. Here there is clear evidence of polarization fields driving external current. The system has resolved to steady state such that the current flows in response to an electrostatic polarization field $\Delta \mathbf{E}$. In such a field,
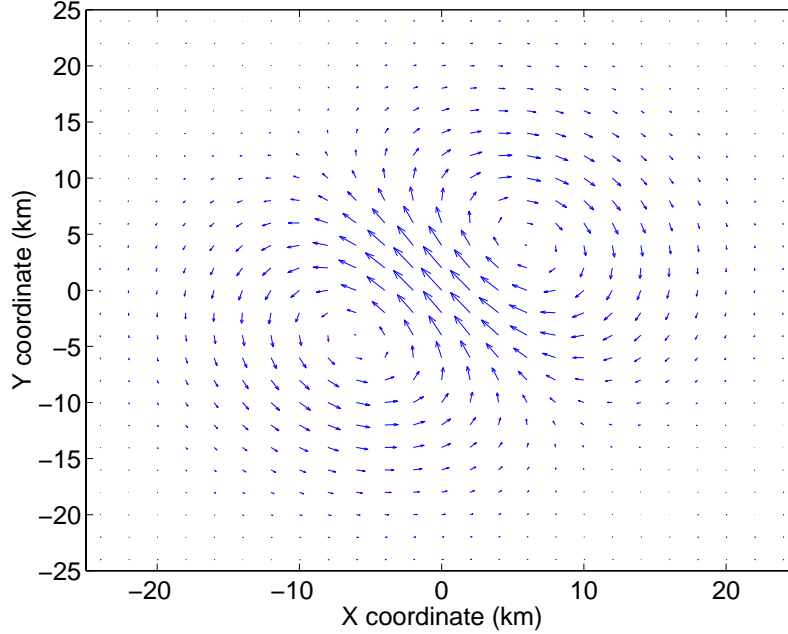
92

Figure 3-25: Current pattern at 2 ms.

the current follows the prediction of Equation (3.64):

$$\phi_0 = \arctan\left(\frac{\Sigma_H}{\Sigma_P}\right) + \alpha. \tag{3.114}$$

For our choice of parameters, ϕ_0 is 126.8 degrees, as is reflected in the direction of the current inside the heated disc. Thus on the long time scale, the perturbed current has contributions from both the non-polarization and polarization components:

$$\delta \mathbf{J} = \delta\boldsymbol{\sigma}\,\mathbf{E}^0 + \overline{\boldsymbol{\sigma}}\,\delta\mathbf{E}. \tag{3.115}$$

In between the two limits, the angle continuously evolves between α and ϕ_0. The direction of the current in the heated volume as a function of time is given in Figure 3-26. Evidently, the angle evolution includes additional 360-degree rotations between the limits. No attempt will be made to interpret the transient response as it does not necessarily reflect the frequency response. However, it does suggest the value of the magnetic diffusion time in the horizontal direction as around 100 μs for this choice of daytime background parameters, which is comparable to the thermal response time
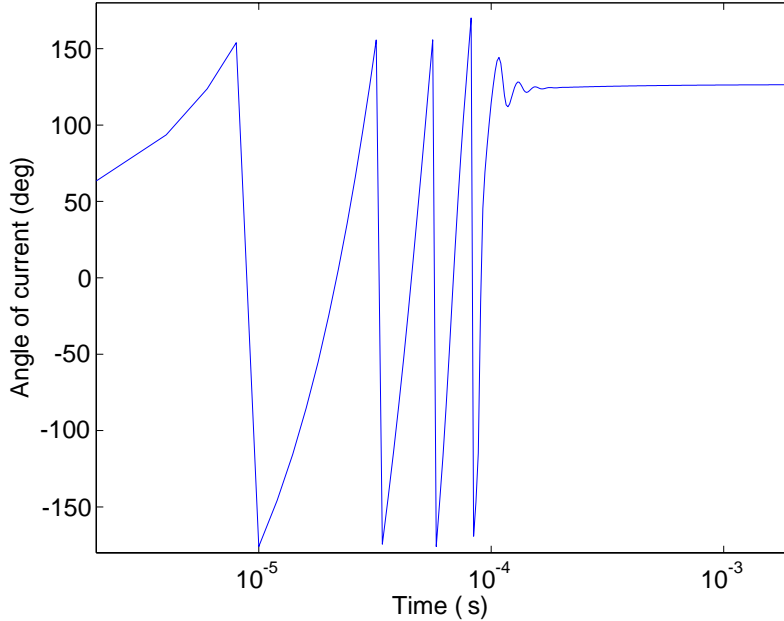
Figure 3-26: Current direction inside heated volume as a function of time.

τ of the plasma. Thus, we do not expect that horizontal loop structure will develop for modulation frequencies much higher than $1/\tau$.

Simulation results have also been obtained in three dimensions. The three dimensional runs are an attempt to reconcile the effects of vertical inhomogeneity and horizontal anisotropy which could be accounted for only crudely in the analytic theory. Again, we consider the fast and slow time scale behaviour, and the location of the transition between them. The simulation was run for a box with dimensions 30 km on a side, discretized into 1-km cubic elements. The daytime conductivity profiles of Figure 3-11 and the 3.3 MHz daytime profiles of fluctuating conductivity from Figure 3-14 were used in the simulation. The response was calculated to a step in the conductivity, with initial fields $\Delta \mathbf{E}$ and $\Delta \mathbf{H}$ assumed to be zero. The time step was chosen to be 20 ns, which was found to be the minimum time step which would allow stability.

The fast time scale behaviour is shown in Figures 3-27 and 3-28. In Figure 3-27 is the J_x and J_z components of the current in the x-z plane, and in Figure 3-28 is the J_y and J_z components of the current in the y-z plane. These are the currents at 100
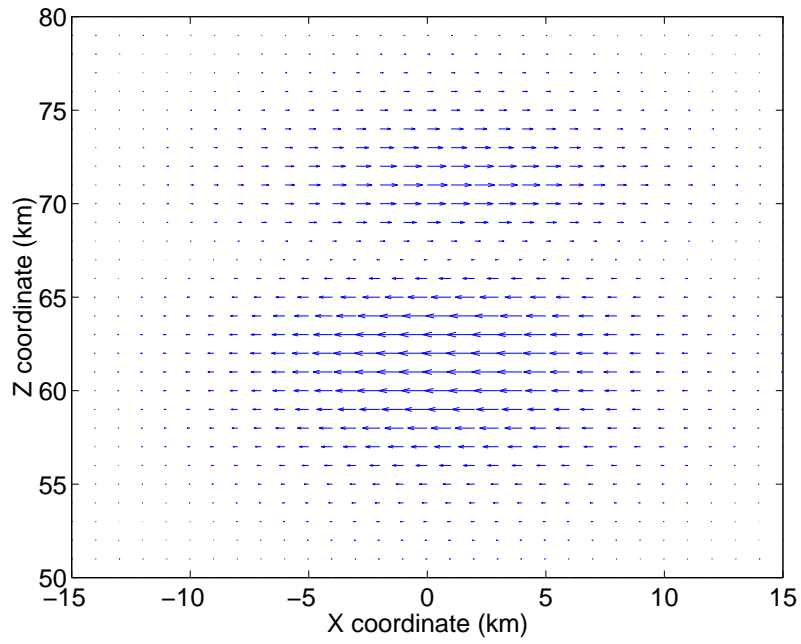
94

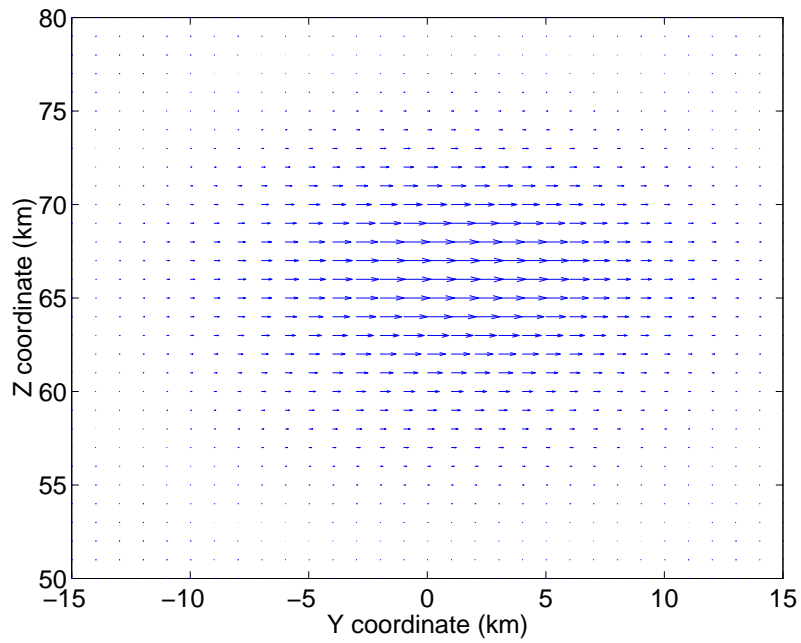Figure 3-27: J_x and J_z in the x-z plane at 100 ns.
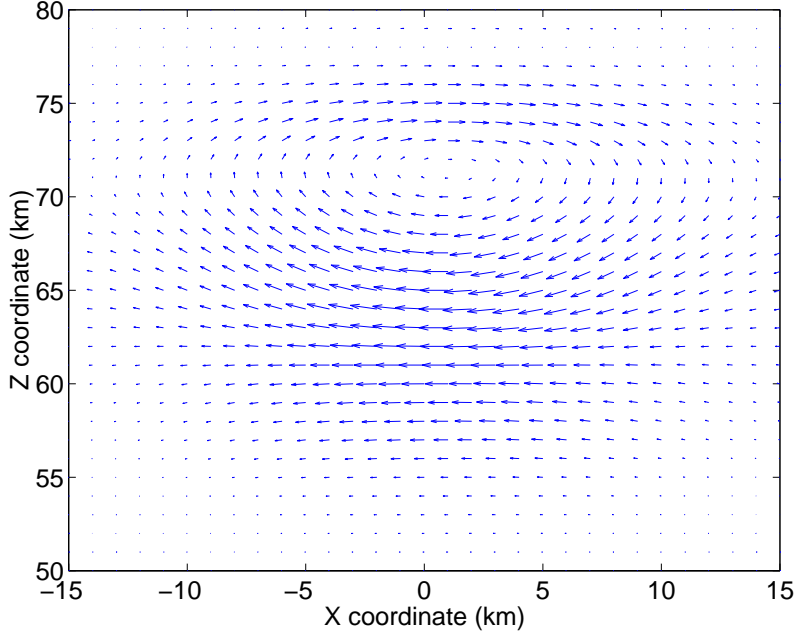

Figure 3-28: J_y and J_z in the y-z plane at 100 ns.

Figure 3-29: J_x and J_z in the x-z plane at 40 μs.

ns of simulation—vastly insufficient time for the development of polarization electric fields.

The perturbation current in this case results entirely from the conductivity modification. Figure 3-27 shows the current $J_x = \delta\sigma_P(x,z)E^0$ and Figure 3-28 shows the current $J_y = -\delta\sigma_H(y,z)E^0$. For example, in Figure 3-27 the positively x-directed current at 72 km and negatively x-directed current at 63 km represents the qualitative features of the Pedersen conductivity modification (Figure 3-14).

On slow time scales, we expect the formation of current loops in the horizontal and vertical directions. The three-dimensional simulation settles into steady state after about 40 μs as opposed to 100 μs in the two-dimensional simulation. The vertical loops evidently form faster than the horizontal loops. Figures 3-29 and 3-30 are the same views as Figures 3-27 and 3-28 except that 40 μs has elapsed since the conductivity step change.

We note the vertical loop structure of the current in both Figures 3-29 and 3-30. The loop is flattened in the vertical direction, suggesting it has a smaller inductance than the horizontal loops above the heated volume (illustrated in Figure 3-21) and
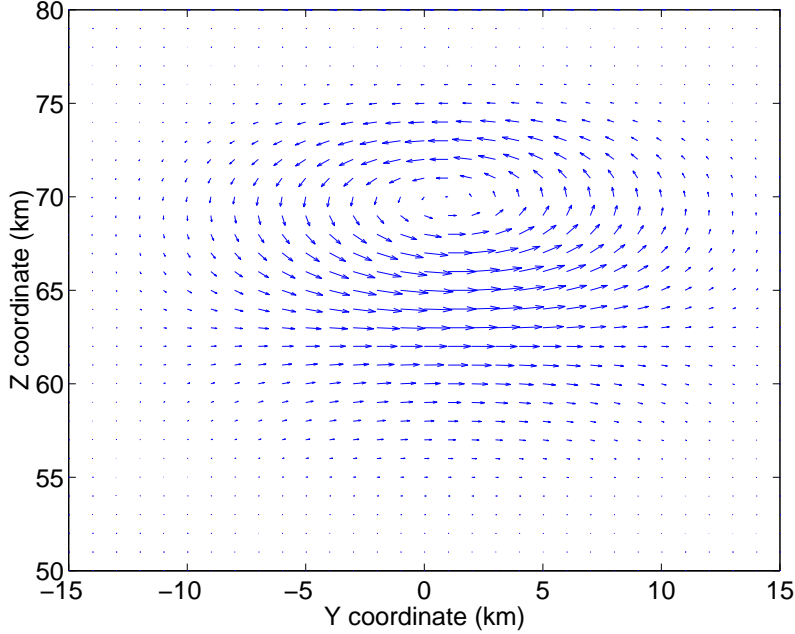
Figure 3-30: J_y and J_z in the y-z plane at 40 μs.

thus is able to form faster. Furthermore, the vertical loop tends to short out the horizontal loops at the altitude of the heated volume (associated with time constant τ'_{mh} in Figure 3-23). Thus the vertical loop dominates all horizontal loops.

In addition, we note that the vertical loop bears none of the structure of the conductivity modification curve. This is the result of the fact that the electric charge relaxation time τ_e in the vertical direction is much faster than the magnetic time constants. Thus the formation of the loops depends on the height-integrated conductivity change rather than the height-dependent conductivity change. Justification for this interpretation can be made by comparing the electric time constant $\tau_e = \epsilon_0/\sigma_o$ and the vertical magnetic time constant $\tau_{mv} = \sigma_0\mu_0 L_h L_v$, where we assign scale sizes $L_h = 20$ km, $L_v = 10$ km. $\tau_e \ll \tau_{mv}$ requires that σ_0 be larger than 1×10^{-7} S/m, which is the case in the region of interest (above 60 km) during the daytime conditions under consideration.

A more graphic justification for the dependence on height-integrated conductivity change can be made by considering the variation of the current profiles J_x or J_y with time after the step. For example, J_x is plotted as a function of altitude and time after
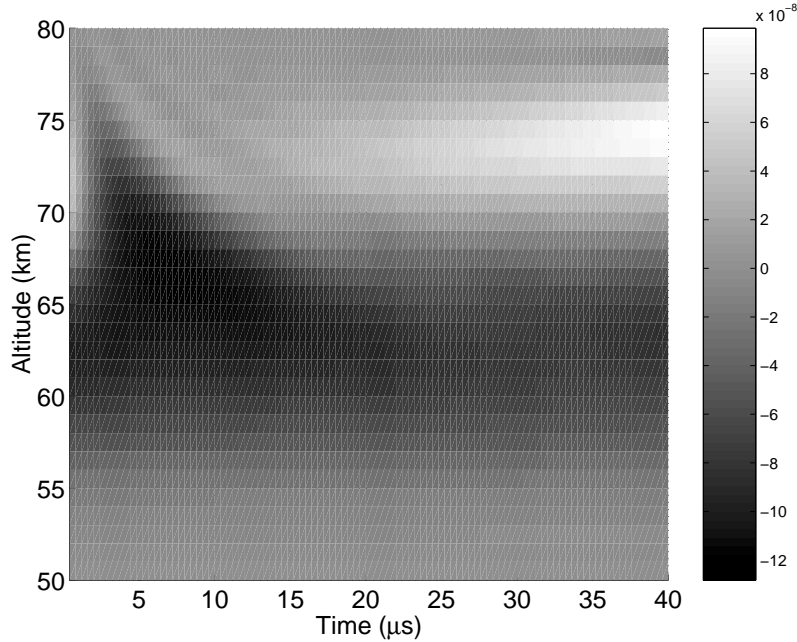
97

Figure 3-31: J_x (grayscale) versus time and altitude.

the step turn-on in Figure 3-31. All structure associated with the height-dependent conductivity change is lost during the first 5 μs, which represents the time τ_e required for charge to relax in the vertical direction. This occurs long before the development of the vertical loop starting at $\tau_{mv} = 15$ μs.

The last point to be made with the three-dimensional simulation is the orientation of the current within the horizontal plane. For simplicity, we consider only the current direction at a location $x = y = 0$ km, with z allowed to vary. The vector $[J_x(z), J_y(z)]$ can be plotted in the horizontal plane (Figure 3-32) as z varies between 50 and 80 km. Here we show the set $[J_x(z), J_y(z)]$ at various time scales. At short time scales (100 ns), the pattern of currents follows the direction of the conductivity modification. The vector rotates clockwise as the Pedersen conductivity modification changes from negative at low altitudes to positive at higher altitudes. At longer time scales, the pattern evolves into a linear structure, rotated about 20 degrees counterclockwise from its initial direction. Some comparison can be made with the predictions of the analytic theory. The height-integrated conductivities for the normal daytime profile
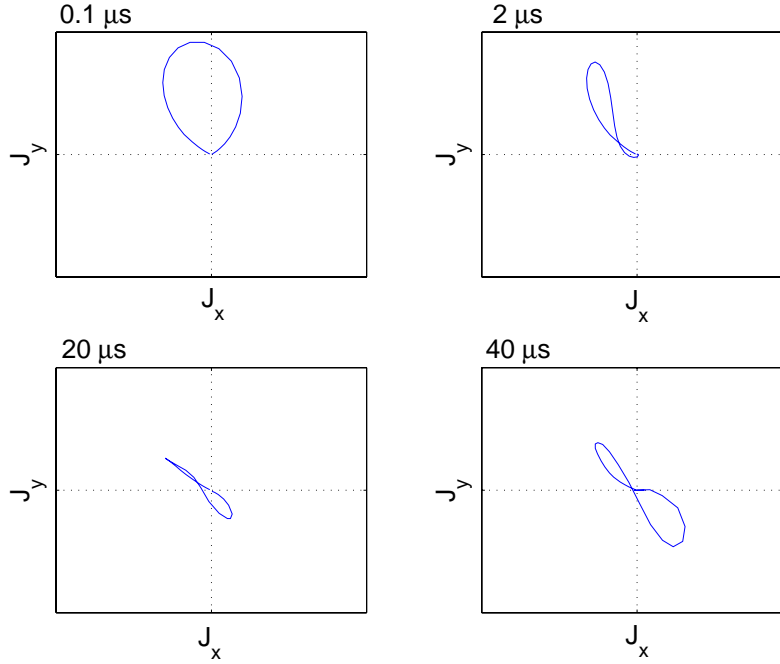
Figure 3-32: Current vector $[J_x(z), J_y(z)]$, z varying between 50 and 80 km.

are

$$\Sigma_P = 3.7 \times 10^{-2} \text{ S} \tag{3.116}$$

$$\Sigma_H = 2.4 \times 10^{-1} \text{ S} \tag{3.117}$$

$$\delta\Sigma_P = -2.3 \times 10^{-5} \text{ S} \tag{3.118}$$

$$\delta\Sigma_H = -1.6 \times 10^{-4} \text{ S.} \tag{3.119}$$

With these parameters, we expect the fast time scale current pattern to be rotated an average (height-integrated) angle of

$$\alpha = -\arctan\left(\frac{\delta\Sigma_H}{\delta\Sigma_P}\right) \tag{3.120}$$

$$= 98 \text{ deg} \tag{3.121}$$

with respect to the $\hat{\mathbf{x}}$ direction, a value which is well-reflected by the first panel of Figure 3-32. At longer time scales, we would expect according to the two-dimensional

theory of Section 3.3.2 that the polarization electric field would rotate the structure so that it is at an angle

$$
\begin{aligned}
\phi_0 &= \arctan\left(\frac{\Sigma_H}{\Sigma_P}\right) + \arctan\left(\frac{\delta\Sigma_H}{\delta\Sigma_P}\right) & (3.122)\\
&= 179 \text{ deg} & (3.123)
\end{aligned}
$$

with respect to the $\hat{\mathbf{x}}$ direction. However, if currents circulating above the heated volume are the dominant structure, then the author's three-dimensional extension (Section 3.3.3) is more applicable, and we expect the current loop to be rotated at an angle close to $\alpha = 98$ deg. The simulation suggests a compromise between these two extremes, with behaviour somewhat closer to the three-dimensional theory. Therefore the suggestion from the simulation is that the horizontal loop structure is weak in comparison to the vertical loop structure after the simulation has settled to steady state. This interpretation can be confirmed by looking at the horizontal current structure in the lower and upper layers of the vertical loop. This is shown in Figures 3-33 and 3-34. These plots are of the horizontal current structure in the lower part of the loop (65 km) and the upper part (75 km) after 40 μs. The plots show that the horizontal loop structure is indeed relatively undeveloped in comparison to the vertical loop structure.
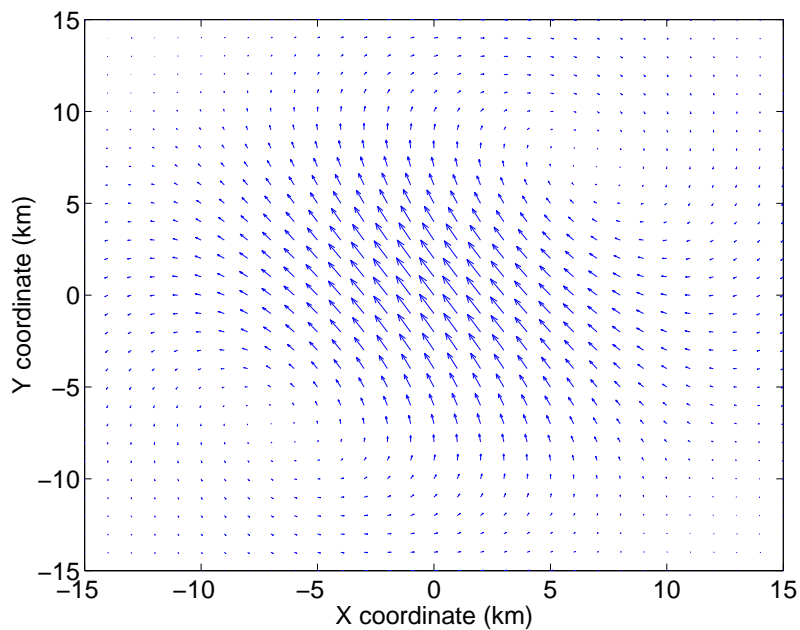
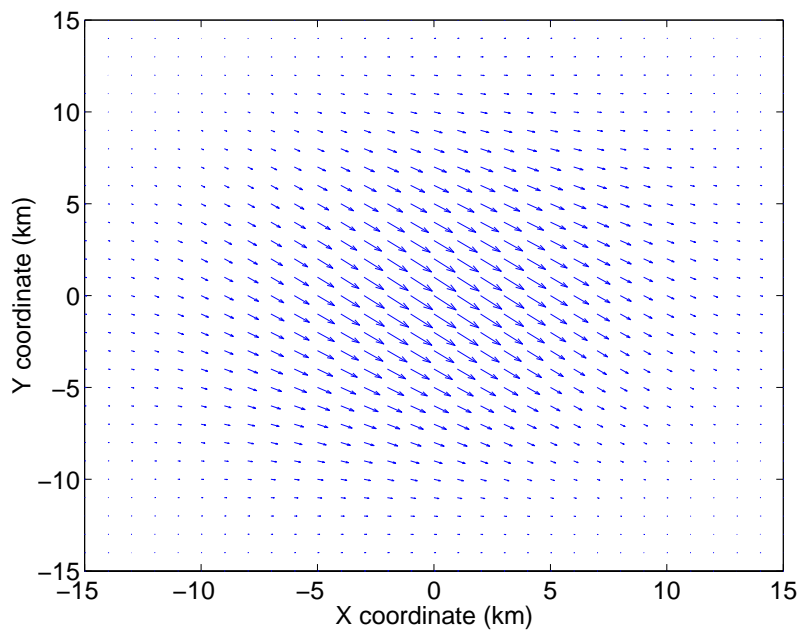Figure 3-33: Horizontal current structure at base of vertical loop (65 km, 40 μs).

Figure 3-34: Horizontal current structure at top of vertical loop (75 km, 40 μs).

Chapter 4

Experiments

Experiments were performed with the HAARP ionospheric heater located at Gakona, Alaska. Reported here are measurements of radiation during ionospheric modification experiments, along with interpretation in terms of the theory presented in Chapter 3.

4.1 Description of the hardware

The HAARP ionospheric heater consists of a CW transmitter and a large antenna array. A photograph of the HAARP antenna array is provided in Figure 4-1. The transmitter produces a maximum power of 960 kW and can be amplitude-modulated in any manner up to 200 kHz. The antenna array consists of 48 crossed-dipole antennas, centered at co-ordinates 62°23'33" N 145°08'48" W. The antennas have tuning gear for the range 2.8–10 MHz, although in practice only a couple small bands within this range are available due to the federal regulation of the radio spectrum. The gain of the array is 14 dB at 3 MHz, increasing to 20 dB at 10 MHz. In addition, the elements can be phased to allow beam steering up to 30 degrees from zenith with a response time of 5 μs, although all experiments described hereafter use a zenith beam.

HF and VLF receiver systems were designed, built and deployed in Alaska by the author. The receiver systems consist of antennas, low-noise amplifiers, filters, and laptop computer-based data acquisition. The HF equipment resided in a trailer

Figure 4-1: The HAARP antenna array.

approximately 50 m from the edge of the HF transmitting array. Its purpose was to record the modulation of the heater during transmission. The VLF receiver system was placed in a trailer at coordinates 62°21'34" N, 145°21'16" W, which translates to approximately 11.4 km from the transmitting array, in an approximate west-southwest direction. The separation from the transmitting array was necessary to ensure that HF signals do not enter the VLF receiver and self-demodulate. The HF and VLF receiver systems were synchronized using the 1 pulse-per-second output of a pair of commercial GPS receivers, which are generally accurate to within $1\,\mu$s.

A block diagram of the HF receiver system is shown in Figure 4-2. In the HF receiver system, fields are sensed with an electric dipole antenna and fed into a trailer by coaxial cable. The electronics in the trailer are shown in Figure 4-3. The antenna signal is sent straight into a crystal power detector. The fields are sufficiently strong next to the transmitting array such that no amplification is necessary. The power detector outputs a signal which is essentially $p(t)$ as defined by Equation (2.140). Superimposed on the data are 20 ms-long pulse-per-second GPS timing pulses. Therefore 980 ms of data are available every second. The combined data and timing pulse signal is sampled continuously at 200 kHz with a 12-bit A/D card in a laptop computer. The
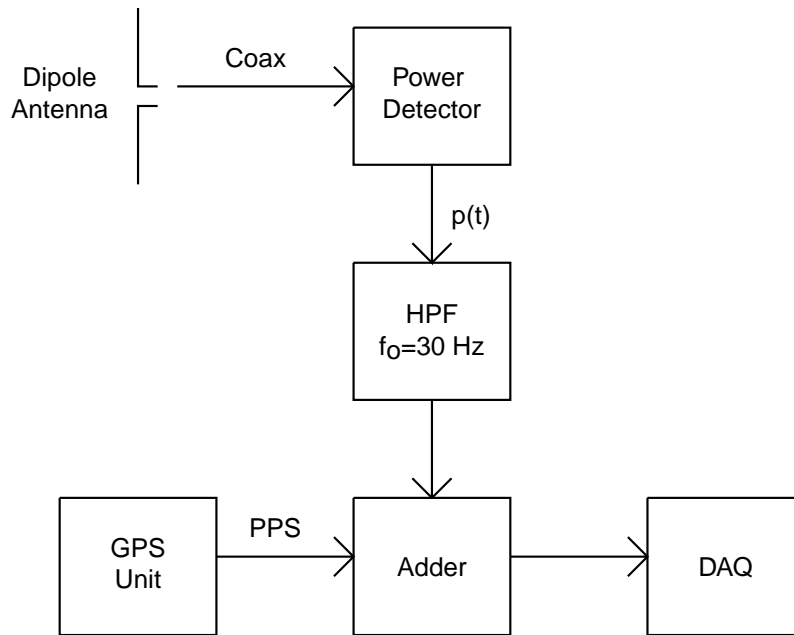
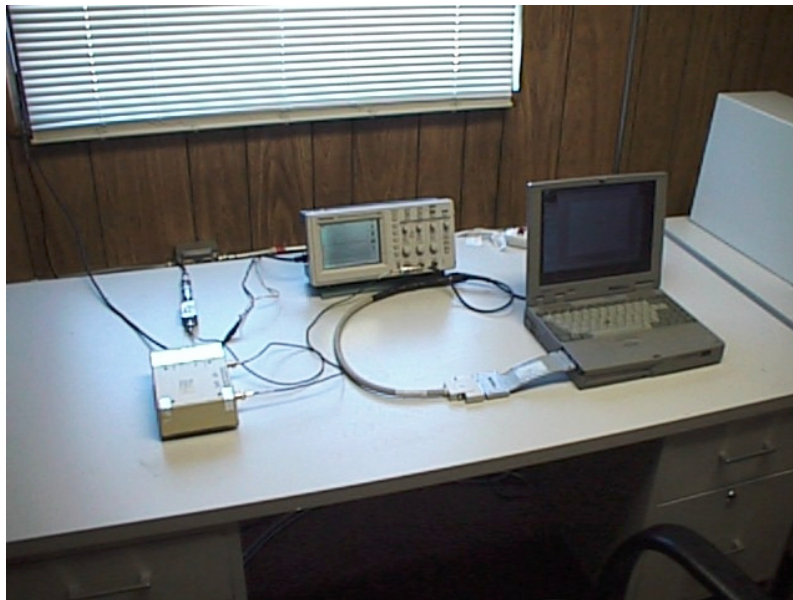Figure 4-2: HF receiver system block diagram.

Figure 4-3: HF receiving equipment.
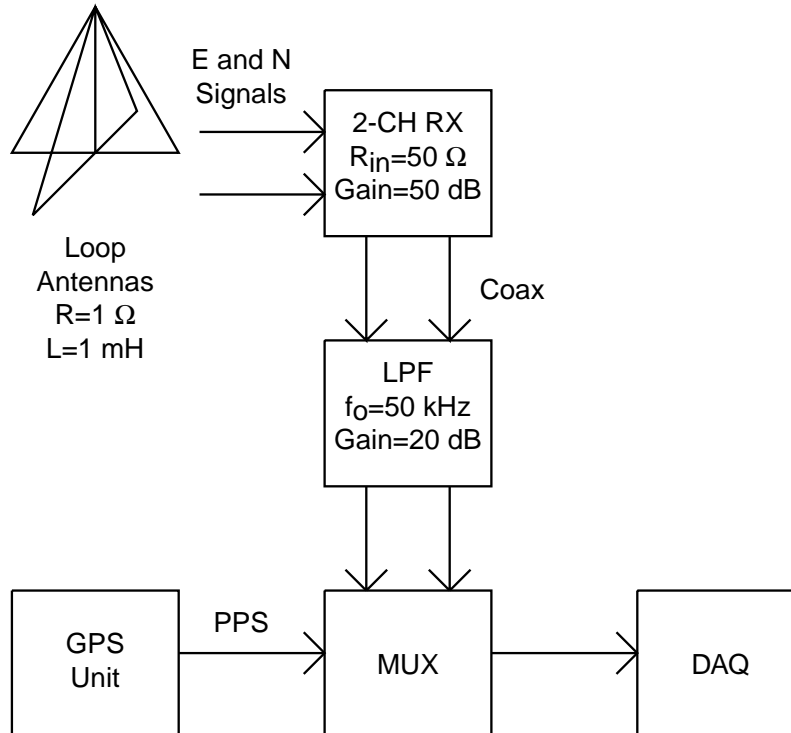
Figure 4-4: VLF receiver system block diagram.

hard disk has a capacity of 36 GB, allowing for 25 hours of data before backup.

A block diagram of the VLF receiver system is given in Figure 4-4. In the VLF receiver system, **H**-fields are picked up with two orthogonal loop antennas (aligned magnetic north-south and east-west), as shown in Figure 4-5. The loops are 5.5 m high with a base of 7.3 m, giving an area of 20 m^2. They are constructed from 6 turns of 14 AWG wire, with a DC resistance of 1.0 Ω and an inductance of 1.0 mH. The receiver is located at the base of the loops and has two input channels, one for each loop. The channels have an input impedance of 50 Ω and a gain of 50 dB. The input coupling is through a step-up transformer into a differential common-emitter pair. The transformer step-up is standard practice to ensure adequate noise performance in the presence of a low-source impedance loop antenna. The voltage divider combination of the loop impedance and the amplifier input impedance results in a receiver signal proportional to $\frac{dH}{dt}$ below 8 kHz, and proportional to H above 8 kHz. The placement of this pole in the receiver response is to de-emphasize 60

Figure 4-5: VLF loop antennas (with author).

Figure 4-6: VLF electronics.

Hz power line harmonics below 8 kHz which would otherwise saturate the limited dynamic range of the 12-bit digitizers. The outputs of the receiver are fed into electronics located in the trailer, shown in Figure 4-6. The signals are lowpass-filtered at 50 kHz, amplified by 20 dB, and fed into a multiplexer. The filter phase response amounts to a 10-μs time delay over the entire band. At 50 kHz, this time delay amounts to a phase of 180 degrees, as one would expect of a four-pole low pass filter. Since only one digitizer channel is available with the data acquisition card, the two antenna signals are alternated every second using the GPS pulse-per-second signal. There is a 20-ms long pulse at every intersignal boundary to provide a precise rising edge for synchronization. The polarity of this pulse is alternated every second so that one will know which antenna signal is being displayed in the 980 ms of data following the pulse.

4.2   Experimental procedure

The experiment consisted of recording the ELF/VLF radiation produced by a modulation frequency sweep of the heater. The modulation frequency sweep was repeated

Date	Time	Carrier	Mod
080602	040000	3.3	90-7170
080602	042000	3.3	90-7170
080602	044000	3.3	90-7170
080702	060000	3.3	90-7170
080702	062030	3.3	90-7170
080702	064100	3.3	90-7170
110902	040000	3.3	85-8365
110902	043000	5.8	85-8365
110902	060000	3.3	85-8365
110902	063000	5.8	85-8365
110902	080000	3.3	85-8365
110902	083000	5.8	85-8365
111002	040000	3.3	85-8365
111002	043000	5.8	85-8365
111002	060000	3.3	85-8365
111002	063000	5.8	85-8365
111002	080000	3.3	85-8365
111002	083000	5.8	85-8365

Table 4.1: Experiment summary.

a total of 18 times. In the first 6 trials (performed 6–7 August 2002), the sweep was between 90 and 7170 Hz in 120 Hz steps (a total of 60 frequencies), with a 10-second dwell on each frequency. After each 10-second dwell, a 10-second "calibration" pulse at 7290 Hz was transmitted. Thus it took 20 minutes to run the entire sweep. In the latter 12 trials (performed 9–10 November 2002), the sweep was between 85 Hz and 8365 Hz in 120 Hz steps (a total of 70 frequencies), with 50-second dwells on frequencies in the range 85 Hz to 1165 Hz and 10-second dwells on frequencies in the range 1285 to 8365 Hz. 10-second calibrations at 7285 Hz were made after each dwell. Thus each sweep was 30 minutes long. The reason for extending the dwell times below 1165 Hz was to improve the detection of signals in that range. The experimental runs are summarized in Table 4.1. Listed are the dates, times, carrier frequencies, and modulation frequencies. The dates and times are given in universal time, which is 8 hours ahead of local time in the first 6 trials, and 9 hours ahead in the remaining 12 trials. The carrier frequency is in MHz, and the modulation frequency in Hz.

4.3 Data processing

The experimental data collected during the trials consist of two time series $s_E(t)$ and $s_N(t)$ corresponding to fluctuations in the east-west and north-south magnetic fields. These directions are determined by the physical orientation of the loop antennas, which stay fixed during the duration of the experiments. These will be the same directions as the theoretical $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ axes of Chapter 3 if $\mathbf{E}^0$ points eastward. However, $\mathbf{E}^0$ in general can point in any direction, thus the physical axes and the theoretical axes differ by an azimuthal rotation around the origin.

We now consider the signal of interest, namely the magnetic field fluctuations at the HF heater modulation frequency. The signal is narrowband, and as such is modelled as a sine wave with a slowly varying magnitude and phase. Thus there are four observables: the magnitudes (A_E, A_N) and phases (ψ_E, ψ_N) of the east-west and north-south magnetic field fluctuations at the modulation frequency.

The phase reference for the ELF/VLF signals is the time series $s_0(t)$ of the modulation leaving the HF transmitter. The total measured phase at the ELF/VLF recording site will therefore consist of phase accrued during propagation from the ground to the ionosphere, during the excitation of radiation in the ionosphere, and during the return to the ground location of the VLF receiver. Any additional phase due to the receiver input impedance and lowpass filter response is removed from the data during the analysis.

If we continue to assume an $e^{-i\omega t}$ dependence of all signals, then a negative phase will correspond to a leading signal and a positive phase will correspond to a lagging signal. With this convention, the signals $s_0(t)$, $s_E(t)$, and $s_N(t)$ have the form

$$s_0(t) = \Re\left(e^{-i\omega t}\right) = \cos(\omega t) \tag{4.1}$$

$$s_E(t) = \Re\left(A_E e^{i\psi_E} e^{-i\omega t}\right) = A_E \cos(\omega t - \psi_E) \tag{4.2}$$

$$s_N(t) = \Re\left(A_N e^{i\psi_N} e^{-i\omega t}\right) = A_N \cos(\omega t - \psi_N). \tag{4.3}$$

If $\hat{\mathbf{x}}$ points east and $\hat{\mathbf{y}}$ points north, then the complex amplitudes of the magnetic

field components are given by

$$\mathbf{H}_E = \hat{\mathbf{x}} A_E e^{i\psi_E} \tag{4.4}$$

$$\mathbf{H}_N = \hat{\mathbf{y}} A_N e^{i\psi_N}. \tag{4.5}$$

For some of the analysis, it is useful to recast the fields in terms of rotating coordinates:

$$\mathbf{H}_R = (\hat{\mathbf{x}} - i\hat{\mathbf{y}}) A_R e^{i\psi_R} \tag{4.6}$$

$$\mathbf{H}_L = (\hat{\mathbf{x}} + i\hat{\mathbf{y}}) A_L e^{i\psi_L}. \tag{4.7}$$

Since the direction of propagation is downward, when viewed from above, the $\mathbf{H}_R$ fields rotate clockwise and the $\mathbf{H}_L$ fields rotate counterclockwise. The transformations between coordinate systems is the usual one:

$$A_R e^{i\psi_R} = \frac{A_E e^{i\psi_E} + i A_N e^{i\psi_N}}{2} \tag{4.8}$$

$$A_L e^{i\psi_L} = \frac{A_E e^{i\psi_E} - i A_N e^{i\psi_N}}{2}. \tag{4.9}$$

The polarization ellipse shown in Figure 4-7 can be expressed easily in terms of the rotating amplitudes and phases. The ellipse parameters are the ellipticity ϵ, orientation angle β, size A, and phase ψ. The ellipticity will be defined as the ratio $\epsilon = a/b$ of the semiminor and semimajor axes, such that $\epsilon = 1$ is right-hand, and $\epsilon = -1$ left-hand circular polarization. The orientation angle β generally takes on values in the range $-\pi/2 < \beta < \pi/2$, due to the π-rotational symmetry of the ellipse. The ellipse size A will be taken to be the absolute length of the semimajor axis b. The phase ψ is the value of ωt at the time when the magnetic field is aligned with the semimajor axis b as shown in Figure 4-7. These ellipse parameters are expressed in terms of the rotating amplitudes and phases as follows:

$$\epsilon = \frac{A_R - A_L}{A_R + A_L} \tag{4.10}$$
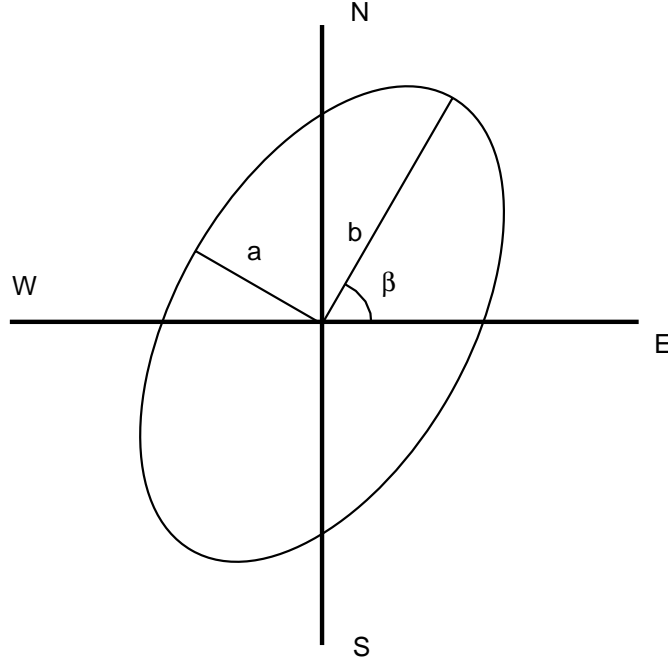
$$\beta = \frac{\psi_R - \psi_L}{2} \tag{4.11}$$

111

Figure 4-7: Polarization ellipse.

$$A = A_R + A_L \tag{4.12}$$

$$\psi = \frac{\psi_R + \psi_L}{2}. \tag{4.13}$$

These parameters provide a complete description of the horizontal magnetic field at the ground. In fact, the sets $(\epsilon, \beta, A, \psi)$, $(A_R, A_L, \psi_R, \psi_L)$, and $(A_E, A_N, \psi_E, \psi_N)$ all contain the same information. The choice of which to use is a matter of convenience.

In order to draw a connection between the current systems described in Chapter 3 and the measured radiation on the ground, we consider the magnetic field to arise from a vector potential:

$$\mathbf{H} = \nabla \times \int \frac{\mathbf{J}(\mathbf{r}')e^{ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \, d\mathbf{r}'. \tag{4.14}$$

If the size of the source region is small compared to the distance to the observer, then

$$\mathbf{H} = \nabla \times \frac{e^{ikr}}{4\pi r} \int \mathbf{J}(\mathbf{r}')e^{-i\mathbf{k}\cdot\mathbf{r}'} \, d\mathbf{r}'. \tag{4.15}$$

If we further assume that the observer is in the far field (the near/far field transition occurs at 650 Hz, at 70 km range), and ignore phase effects from non-zenith contributions to the radiation, then

$$\mathbf{H} = \frac{i\mathbf{k}Ae^{ikz}}{4\pi z} \times \int \mathbf{J}(z)e^{-ikz}\,dz. \tag{4.16}$$

In particular, the current outside the heated volume is dominantly dipolar, so its contribution to the vector potential vanishes in the azimuthal integration of Equation (4.15). Thus the area A of Equation (4.16) refers to the cross-sectional area of the heated volume alone. In the fast modulation limit (modulation frequency above approximately 1–2 kHz), the current is proportional to the integral of the heating function $p(t)$ [Eq. (3.13)]. Thus in frequency domain, the fast modulation conductivity profiles and currents are the values given in Chapter 3 [such as Figure 3-14 and Equation (3.36)], multiplied by the imaginary unit i. $\mathbf{J}$ in Equation (4.16) above can be written $i\mathbf{J}_c$ where $\mathbf{J}_c$ is any calculated current profile from Chapter 3 in the fast modulation limit. If we break $\mathbf{J}_c$ into components, then we have that

$$\mathbf{H} = \frac{kAe^{ikz}}{4\pi z}\left[-\hat{\mathbf{x}}\tilde{J}_{cy}(k) + \hat{\mathbf{y}}\tilde{J}_{cx}(k)\right], \tag{4.17}$$

where the tilde notation refers to a Fourier transform. Thus we have the immediate conclusion that the k-spectrum of the radiation is the Fourier transform of the fluctuating current profile in the fast modulation limit. For example, an impulse-shaped current profile $\delta(z)$ would give rise to a flat radiation spectrum, with zero phase. However, a doublet-shaped distribution $\dot{\delta}(z)$, would also give a flat spectrum, but with 90 degrees of leading (negative) phase.

Furthermore, since the quantities J_{cx} and J_{cy} are themselves real, we would also expect from Equation (4.17) that the radiation be linearly polarized, rotated 90 degrees clockwise (as viewed from above) to the direction of the fluctuating current.

With this framework in mind, we can examine the radiation. We will start with ellipticity and ellipse orientation, and then move to the magnitude and phase. The latter two quantities will be inverse Fourier-transformed to yield an experimentally

113

determined current profile along the lines of Equation (4.17).

4.3.1 Measurements of ellipticity

As just discussed, we expect the radiated fields to be linearly polarized, at some angle with respect to $\mathbf{E}^0$. However, the lower ionospheric plasma supports only right-handed propagation. Thus we might expect that the L-component of a wave attenuates relative to the R-component, resulting in right-hand elliptical polarization at the ground. Let us consider the parallel refractive index [Eq. (3.14)] for ELF/VLF radiation at the modulation frequency ω_m, with $\omega_m \ll \nu_e, \omega_c$:

$$n_\parallel^2 = 1 \pm \frac{\omega_p^2(\omega_c \pm i\nu_e)}{\omega_m(\omega_c^2 + \nu_e^2)}. \tag{4.18}$$

If propagation effects (such as L-mode attenuation) actually occur, they will generally manifest themselves at altitudes high enough that we can take $\omega_c \gg \nu_e$. Furthermore, we consider the case where ω_m is small enough such that we can drop the $+1$ term. What remains is the helicon relation:

$$n_\parallel^2 = \pm\frac{\omega_p^2}{\omega_m\omega_c}. \tag{4.19}$$

The size of the L-mode relative to the R-mode is given by the integration of the imaginary refractive index over the propagation path:

$$A_L/A_R = \exp\left[-\frac{\omega_m}{c}\int \Im\left(n_{\parallel L}\right)\,dz\right]. \tag{4.20}$$

A_L/A_R is related to the Faraday rotation ψ_F since $\Im\left(n_{\parallel L}\right) = \Re\left(n_{\parallel R}\right)$ for helicons:

$$
\begin{aligned}
\psi_F &= \frac{\omega_m}{c}\int \Re\left(n_{\parallel R} - n_{\parallel L}\right)\,dz, & (4.21)\\
&= -\log\frac{A_L}{A_R} & (4.22)\\
&= \log\frac{1+\epsilon}{1-\epsilon}. & (4.23)
\end{aligned}
$$

Ellipticities for the 18 experimental runs are plotted versus modulation frequency in Figure 4-8. On the whole, the ellipticities are close to zero, implying linear polarization and little Faraday rotation effect. The interpretation of these results is that the radiation source is at a low enough altitude that the propagation effects described above can be ignored. For our purposes, we will consider the radiation to occur in free space. In a few cases the ellipticity is significant and negative (L-polarized) and therefore a different propagation model than the above (such as a waveguide model) would be required to properly interpret the data.

4.3.2   Ellipse orientation

We wish to compare the theoretical ellipse orientation angle (where the axes are determined by $\mathbf{E}^0$ pointing along $\hat{\mathbf{x}}$) with the physically measured ellipse angle (where the axes are determined by the antenna placement). The connection between the two sets of axes lies in determining the direction of $\mathbf{E}^0$ during the experiments.

Thus we will first determine $\mathbf{E}^0$. Second, we will estimate the expected theoretical polarization angle α for the currents in the heated volume. Third, we will plot the theory alongside the experimental measurements of the radiated fields.

To determine the direction of $\mathbf{E}^0$, we consult convection maps of the high-altitude auroral plasma flow, similar to the one shown in Figure 1-7. These maps are available through the CEDAR/TIMED Space Weather initiative, and are compiled from a variety of satellite, radar, and modelling sources. From these maps, the electric field vectors can be obtained by rotating the flow vectors clockwise by 90 degress, in accordance with the ideal MHD "Ohm's law" relation $\mathbf{E} \approx -\mathbf{v} \times \mathbf{B}$ in the collisionless upper ionosphere. An example of a convection map (0600 UT, 10 Nov 2002) is shown in Figure 4-9.

In the map, there are contours showing the flow trajectories, as deduced from satellite measurements and modelling, as well as clusters of individual vectors consisting of HF backscatter radar measurements of the flows (note that no radar measurements appear in Figure 1-7). Since the electric field is normal to the flow vectors, the flow trajectories can be thought of as equipotential lines. The contour voltages are labelled
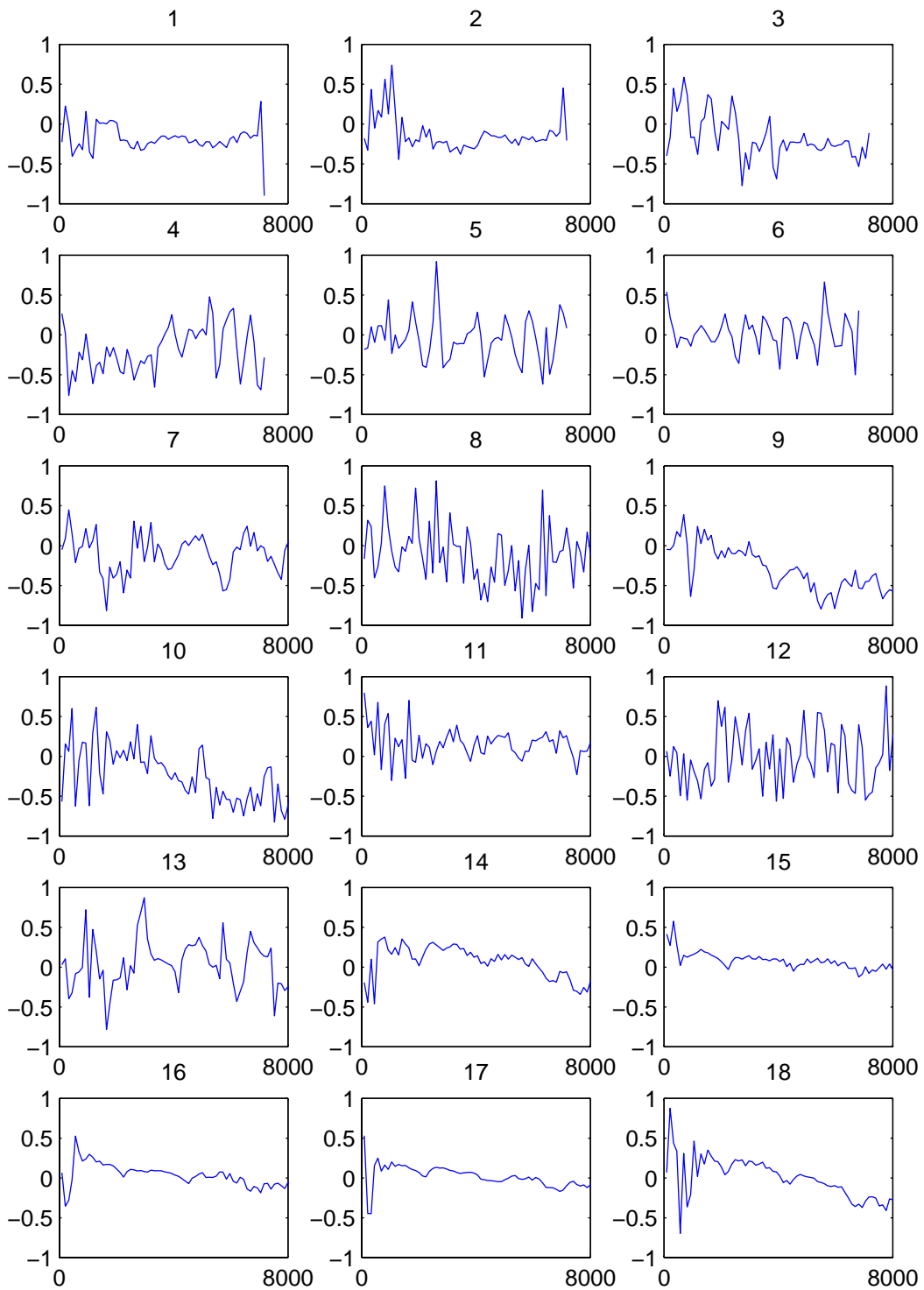
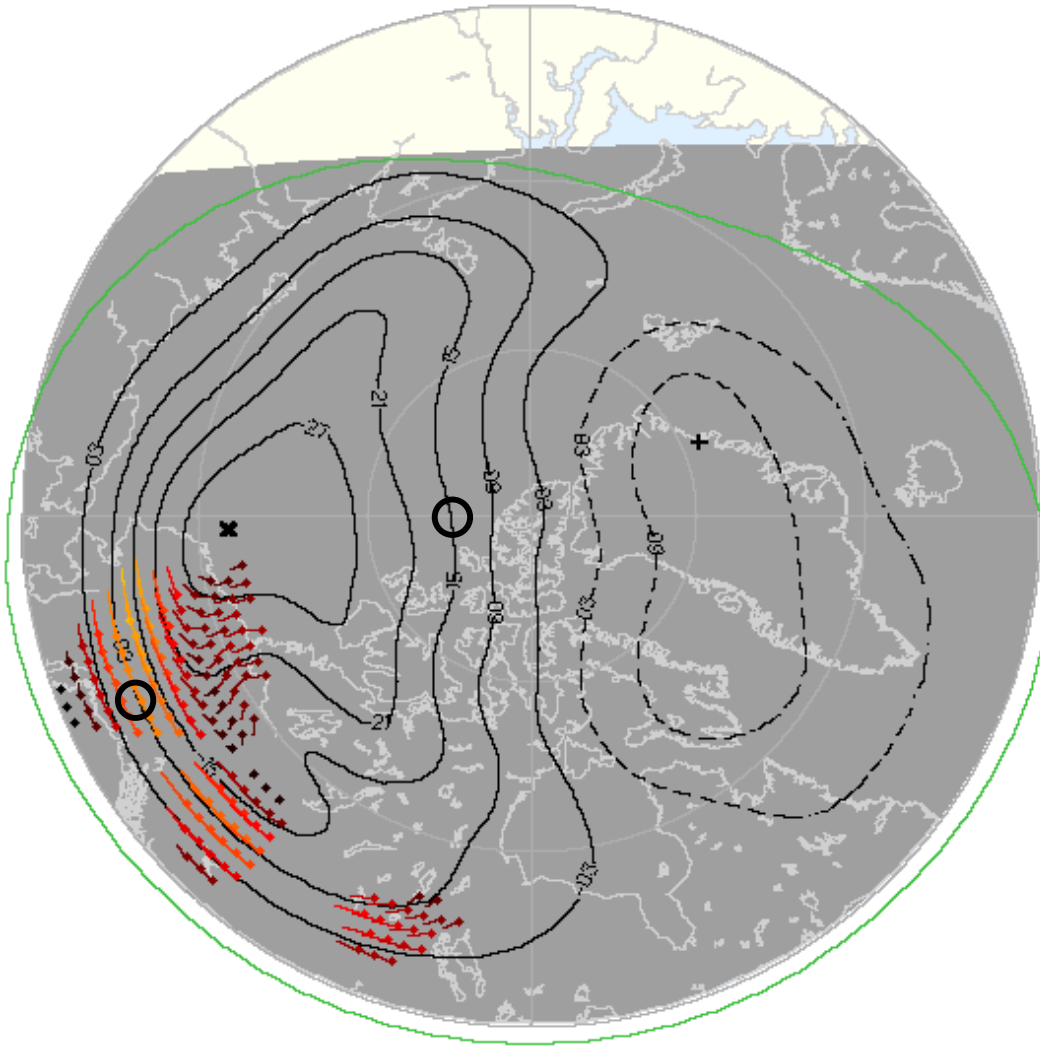Figure 4-8: Ellipticity ϵ versus modulation frequency (Hz).

Figure 4-9: CEDAR/TIMED convection map.

in the diagram, in increments of 3 kV. The HAARP facility is located at approximately the 8 o'clock position in this diagram (denoted by a circle), and the magnetic pole is located at the intersection of the 15 kV contour and the horizontal mid-axis (denoted by another circle). Rotating the velocity vectors over HAARP by 90 degrees yields an electric field direction about 8 degrees clockwise of magnetic north.

A check for the convection map is provided by the magnetometer measurements. The magnetometer senses the three components of the DC magnetic field on the ground due to electrojet current driven by $\mathbf{E}^0$. To deduce $\mathbf{E}^0$ from these DC magnetic fields, some assumption has to be made about the DC conductivity in the overhead plasma. In most cases the Hall conductivity is dominant, and thus the DC magnetic field points in the general direction of $\mathbf{E}^0$. Some caution needs to be exercised as the measured DC magnetic field involves contributions from all currents in the environment, not just the overhead ones. However, the overhead contribution often dominates, and so the magnetometer measurements may be viewed as a confirmation of the direction of $\mathbf{E}^0$ as determined from the convection maps. A magnetometer plot (from the HAARP instrument suite) for 10 Nov 2002 is shown in Figure 4-10. The strong northward component in the magnetometer trace (bottom panel) at 0600 UT suggests an eastward electrojet current, and thus a generally northward electric field $\mathbf{E}^0$, confirming the conclusion from Figure 4-9. Thus we can determine the direction of $\mathbf{E}^0$ for the 18 experimental runs in this manner. This is straightforward for all cases, except for experiment runs 1–3. These three data sets were taken during the daytime, and the convection map pattern was shifted away from the Alaska region during these trials. Therefore the direction of $\mathbf{E}^0$ was taken from magnetometer measurements alone in these cases, and as a result they should not be taken too seriously.

Having determined $\mathbf{E}^0$, the next step is to form the theoretical estimate of the polarization angle. We estimate an angle $\alpha = -\arctan(\delta\Sigma_H/\delta\Sigma_P)$. To find α, we need to height-integrate the profiles of fluctuating conductivity (such as Figure 3-14). In calculating the conductivity, we have to choose an HF carrier frequency and plasma density profile which matches the experiment. The HF carrier frequency is known because it is specified by the operator. However, the plasma density profile will have
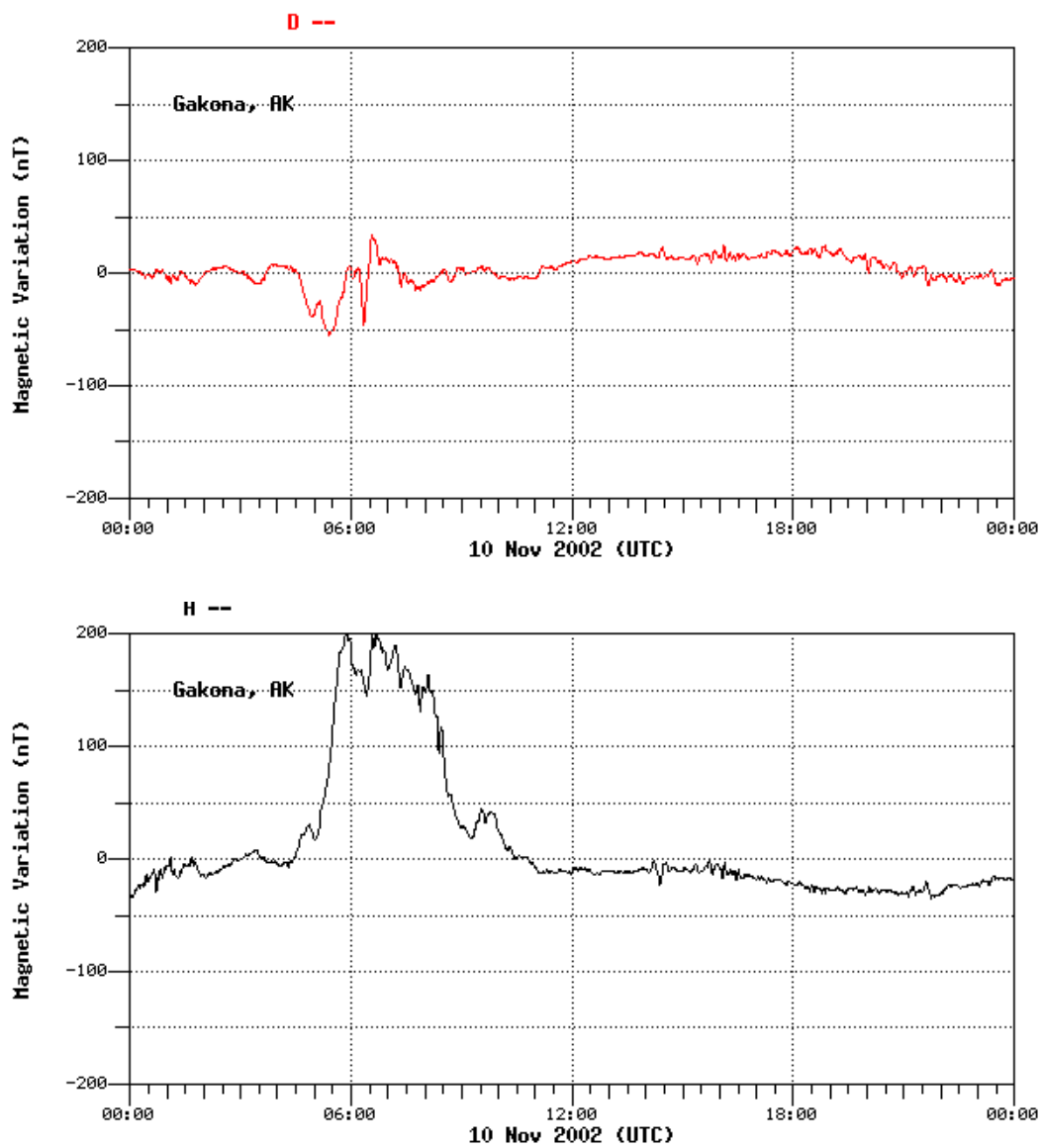
Figure 4-10: Magnetometer data. Top trace is magnetic east component, bottom trace is magnetic north component.

Date	Time	D/N	Carrier	Rio	Density
080602	040000	Day	3.3	0.1	Low
080602	042000	Day	3.3	0.2	Low
080602	044000	Day	3.3	0.4	Low
080702	060000	Night	3.3	0.1	Low
080702	062030	Night	3.3	0.1	Low
080702	064100	Night	3.3	0.1	Low
110902	040000	Night	3.3	0.4	Medium
110902	043000	Night	5.8	0.4	Medium
110902	060000	Night	3.3	0.6	Medium
110902	063000	Night	5.8	0.4	Medium
110902	080000	Night	3.3	0.4	Medium
110902	083000	Night	5.8	0.3	Medium
111002	040000	Night	3.3	0.7	Medium
111002	043000	Night	5.8	0.8	Medium
111002	060000	Night	3.3	2.0	High
111002	063000	Night	5.8	1.5	High
111002	080000	Night	3.3	1.3	High
111002	083000	Night	5.8	1.0	High

Table 4.2: Experimental conditions.

to be chosen from the six profiles presented in Chapter 1 (Figure 1-10).

The 18 data sets of Table 4.1 can be classified according to night/day condition, carrier frequency, and lower ionosphere plasma density (ascertained from the riometer absorption). These conditions are used to calculate conductivity profiles for each experimental trial, along with an associated $\delta\Sigma_P$, $\delta\Sigma_H$, and α. Furthermore, each trial features an estimate for the direction of $\mathbf{E}^0$ during the 20- or 30-minute period of the trial. This information is summarized in Tables 4.2 and 4.3. The riometer measurements are in dB and the density characterization corresponds to the profiles from the catalog (Figure 1-10). The height-integrated conductivities are given in units of μS. The angle α is measured in degrees counterclockwise from $\mathbf{E}^0$ and the angle of $\mathbf{E}^0$ is in degrees counterclockwise from magnetic east. The predicted polarization ellipse orientation angle $\angle\mathbf{H}$ for the radiation is simply $\alpha + 90$ degrees, since $\mathbf{H}$ radiates at 90 degrees to $\mathbf{J}$ in the far field. Note that around 0800 UT the location of HAARP shifts from the evening cell (northward $\mathbf{E}^0$) to the morning cell (southward $\mathbf{E}^0$), and hence we observe a large change in the angle of $\mathbf{E}^0$ at that time.

Date	Time	$\delta\Sigma_P$	$\delta\Sigma_H$	α	$\angle\mathbf{E}^0$	$\angle\mathbf{H}$
080602	040000	27.5	-339	85	22	-163
080602	042000	27.5	-339	85	27	-158
080602	044000	27.5	-339	85	30	-155
080702	060000	11.6	-27.1	67	87	-116
080702	062030	11.6	-27.1	67	90	-113
080702	064100	11.6	-27.1	67	92	-111
110902	040000	50.9	-126	68	83	-119
110902	043000	25.4	-38.2	56	70	-144
110902	060000	50.9	-126	68	81	-121
110902	063000	25.4	-38.2	56	78	-136
110902	080000	50.9	-126	68	79	-123
110902	083000	25.4	-38.2	56	30	176
111002	040000	50.9	-126	68	84	-118
111002	043000	25.4	-38.2	56	90	-124
111002	060000	51.0	-362	82	82	-106
111002	063000	42.3	-123	71	71	-128
111002	080000	51.0	-362	82	-102	70
111002	083000	42.3	-123	71	-95	66

Table 4.3: Determinations of α and $\mathbf{E}^0$.

With the determination of $\mathbf{E}^0$ and α completed, we turn to the experiments. The ellipse orientation as defined in Equation (4.11) takes on values in the range $-\pi/2 < \beta < \pi/2$. However, to compare the observations of the magnetic field vector with the theory we need to know the four-quadrant position of the vector.

To illustrate a four-quadrant determination, let us consider the polarization to be linear. This is nearly always the case, as shown by the ellipticity measurements. Thus the polarization ellipse is a line through the origin at some angle to the east-west axis. As the received ELF/VLF signal oscillates in time, the magnetic field vector jumps between two opposite quadrants of the horizontal plane. For example, in Figure 4-11, the vector would jump between the first and third quadrant every half wave period. If one does not care about the phase of the signal, then it does not matter which of the two quadrants the vector is in at any particular time. However, we are concerned with the phase in this study, so it is necessary to know in which of the two quadrants the vector is located at any point in time.

The simplest way to make this determination is to observe the magnetic field

HN
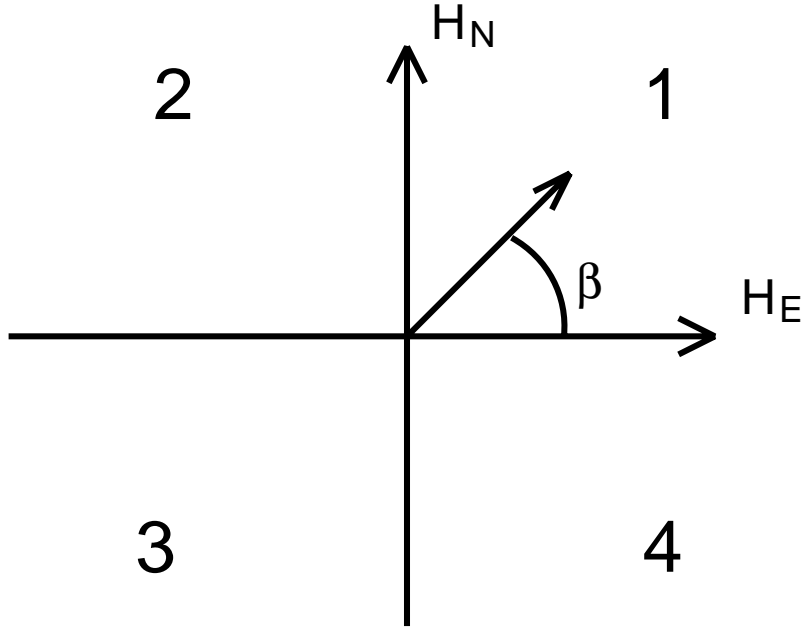
2   1

β   HE

3   4

Figure 4-11: Four quadrants of observation plane.

vector as the modulation frequency is lowered towards zero. In the limit of low frequency, the magnetic field direction is given simply by the right-hand rule for DC current. For example, if the magnetic field vector appears in the first quadrant in the low frequency limit, as shown in Figure 4-11, then the overhead current vector would be located in the fourth quadrant. Furthermore, if we were to look at the computed phase of the signals H_E and H_N, both of them would approach zero degrees in the low frequency limit.

As a second example, let us say the current was in the first quadrant, producing a magnetic field in the second quadrant. In this case, the phase of H_E would approach 180 degrees at low frequency, whereas the H_N phase would approach 0 degrees. A third and fourth case would consist of the low-frequency magnetic field being located in the third and fourth quadrants, respectively.

Assuming that the linear phase factor e^{ikz} dominates the phase of the measured field [Eq. (4.17)], then anticipated phase functions for the four possible cases are shown in Figure 4-12.

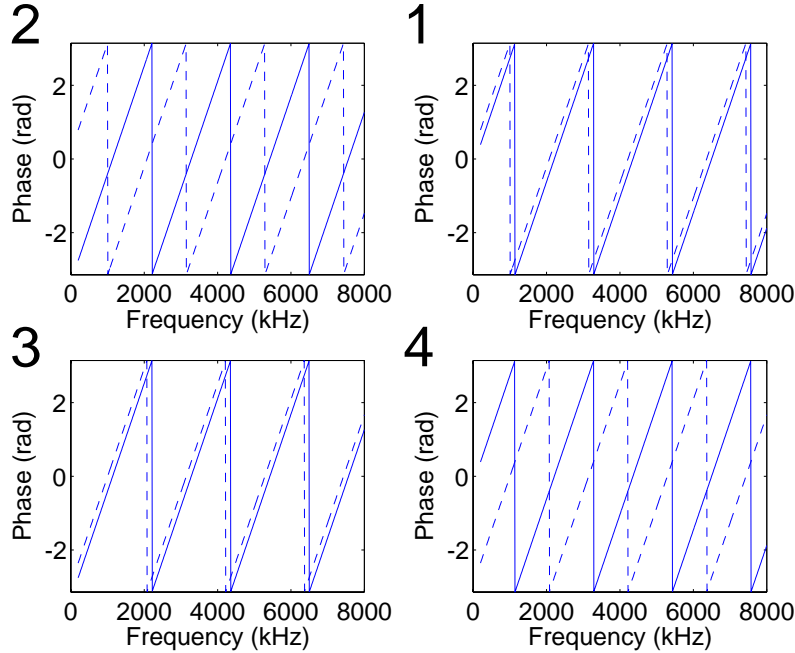We now examine the experimental data. Plotted in Figure 4-13 is the measured

Figure 4-12: Expected measured phase for the H_E and H_N components for the cases of the magnetic field vector in each of the four quadrants, H_E component (solid), H_N component (dashed).

phase of the H_E and H_N components during the 18 experimental runs. Of the data sets, six of them (11,14,15,16,17,18) occur during periods when $\mathbf{E}^0$ is observable, and furthermore these six also exhibit phase relations which clearly fit one of the four cases of Figure 4-12 (all of them being the 4th quadrant case). For these six sets, we can determine the four-quadrant ellipse orientation, and make a comparison with the theory. Of the remaining sets, (1,2) are classifiable as first quadrant, but do not have convection map measurements of $\mathbf{E}^0$. For these sets, we know the four-quadrant ellipse orientation, but have no basis for comparison with the theory. The remaining sets only produce two-quadrant orientation readings, and cannot as such be compared with the theory.

We merge the predictions from Table 4.3 with the experimental measurements of ellipse orientation. Please refer to Figure 4-14. The plots show the orientation angle of the ellipse as a function of modulation frequency. The angular coordinate is the orientation angle, with magnetic north at 12 o'clock and magnetic east at 3 o'clock
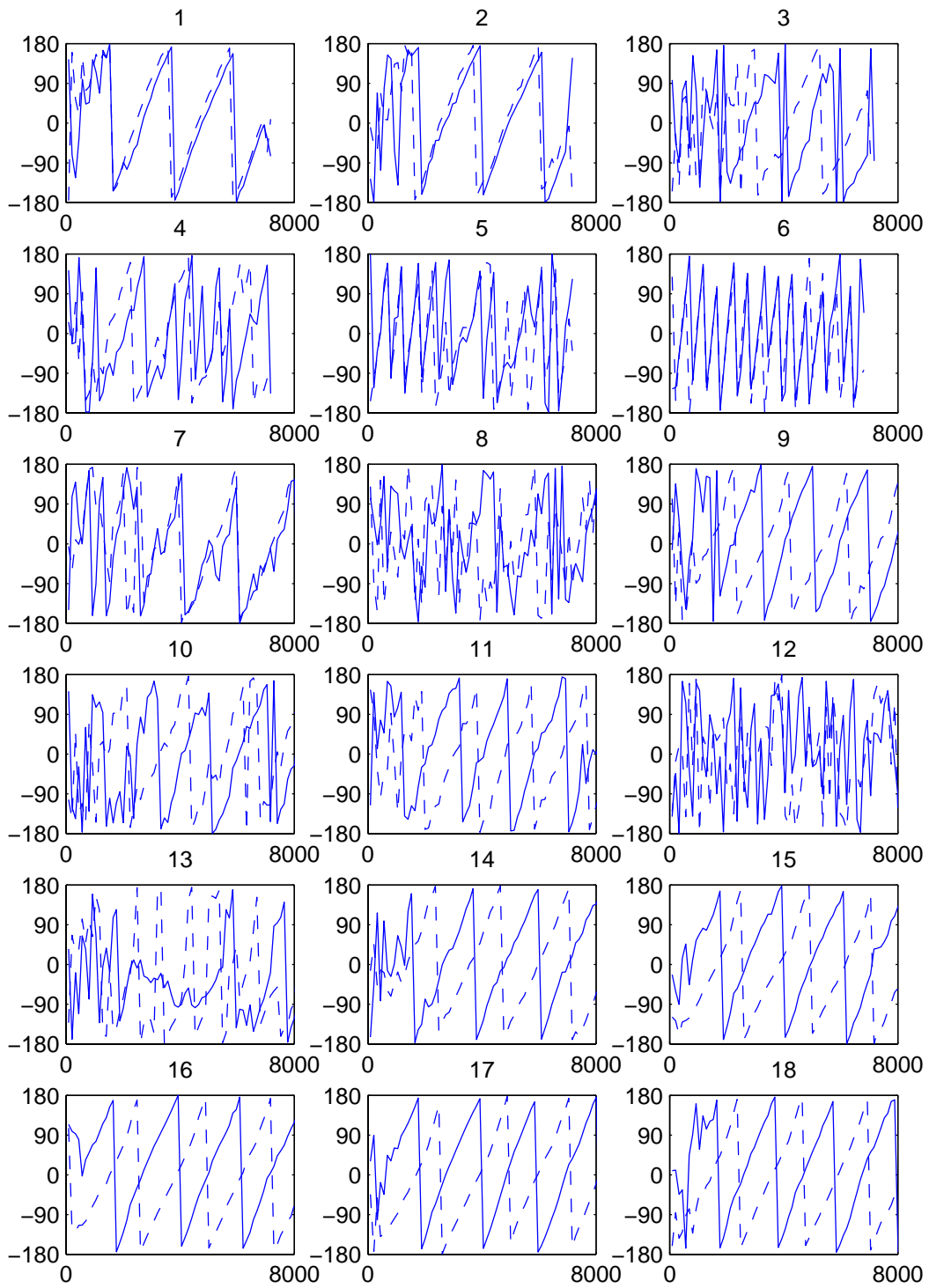
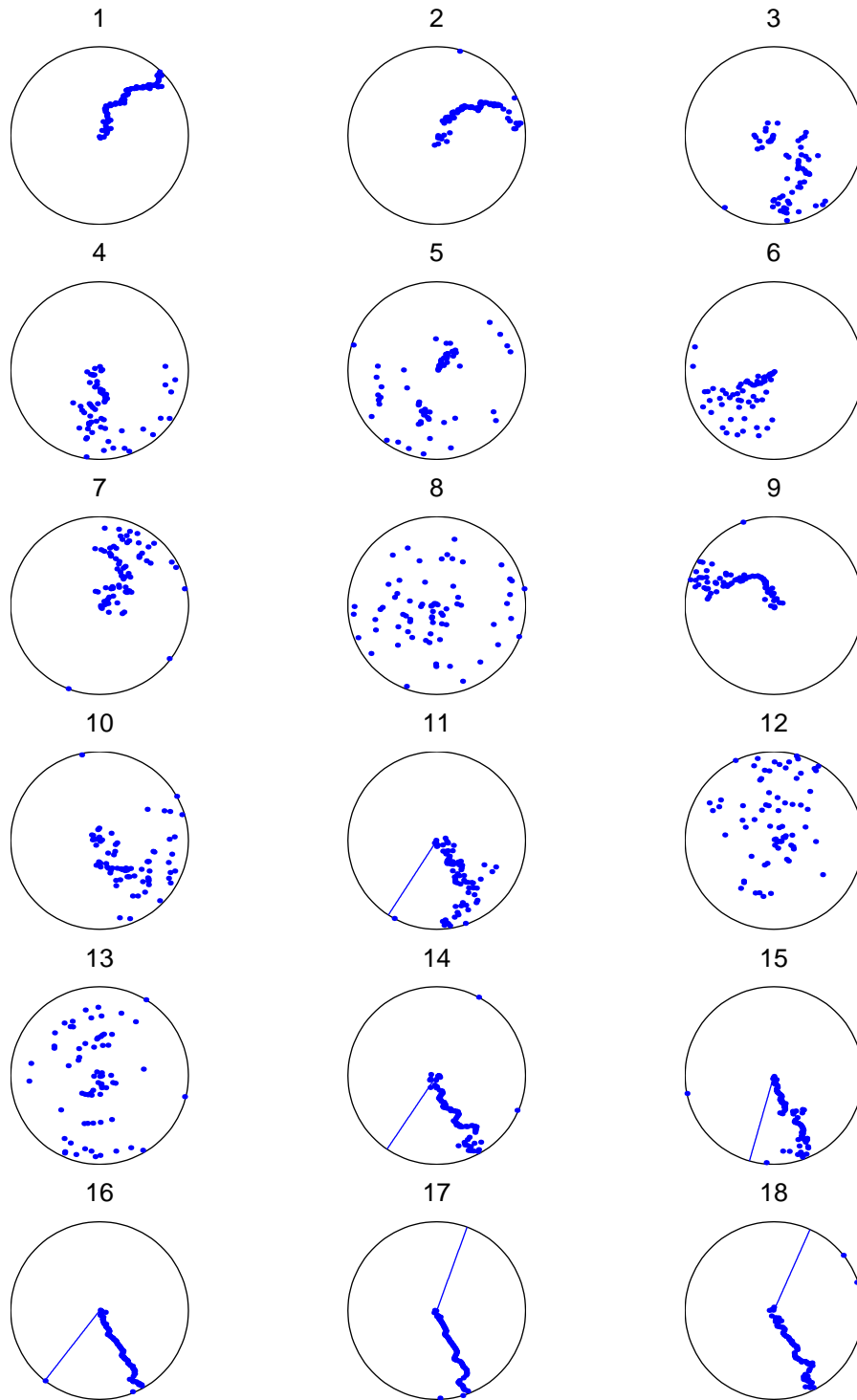Figure 4-13: Measured phase of the H_E component (solid), and H_N component (dashed).

Figure 4-14: Ellipse orientation angle versus modulation frequency, with predictions. Solid line is the predicted angle of **H** if current was directed at angle α with respect to $\mathbf{E}^0$. Dots are experimental measurements.

(labels have been left off of the diagrams to avoid clutter). The orientation angle has been unwrapped as best possible to allow variation over a range of 2π and to facilitate comparison with the theory. The radial coordinate is the modulation frequency, with the origin at zero frequency and the outside of the circle at the maximum modulation frequency. In general, the dependence of orientation angle on modulation frequency is weak, although some plots are scattered and show no pattern at all for the orientation angle, which can be attributed to low signal levels. In plots (11,14,15,16,17,18), the straight line shows the expected direction of the **H** field if the current is oriented at an angle α with respect to $\mathbf{E}^0$. The data of plots (11,14,15,16) show that the actual current is rotated about 30-50 degrees counterclockwise with respect to the calculated α. Plots (17,18) show a clockwise rotation of 130-150 degrees. What is probably happening here is that the satellite/modelling algorithm has prematurely predicted the 180-degree shift in the direction of $\mathbf{E}^0$ which occurs when a location moves from the evening to morning sector. The consistency of the measured radiation in sets (15,16,17,18) suggests the convection map is likely wrong for sets (17,18).

The basic conclusion that will be drawn here is that the magnetic field measurements are rotated approximately 30-50 degrees from α. A similar discrepancy was observed by *Rietveld* [1987], who was searching for evidence supporting the theory of *Stubbe et al.* [1982] (Section 3.3.1). In contrast, if the theory of *Stubbe and Kopka* [1977] is used, the additional rotation of $\arctan(\Sigma_H/\Sigma_P)$ [Eq. (3.64)] is nearly 90 degrees, and thus ϕ_0 would tend to overestimate the rotation of the polarization. With the assumed density profiles, neither theory explains the data. There are at least three possibilities:

1. The assumed density profiles are too weak. A stronger density profile (such as a daytime profile) can give rise to $\delta\Sigma_P < 0$ (see Figure 3-14). This would increase α beyond 90 degrees and could provide the additional rotation such that *Stubbe et al.* [1982] (Section 3.3.1) can explain the data.

2. The theory of *Stubbe et al.* [1982] (Section 3.3.1) is inadequate and the three-dimensional extension of Sections 3.3.3 and 3.3.4 is required. The simulation of

Section 3.3.4 provides additional rotation of the currents beyond α.

3. A combination of the above.

We conclude here that the polarization is ambiguous and cannot alone adequately assess the validity of the theories. Thus we turn to the magnitude and phase to provide further information.

4.3.3 Magnitude and phase

The magnitude and phase of the radiation is controlled by two influences. The first influence is the vertical distribution of current [Eq. (4.17)]. The radiation is the Fourier transform of the vertical profile of the antenna current, multiplied by a linear phase factor e^{ikz}.

The second influence is the resonance of the earth-ionosphere cavity. During modulated electrojet heating, the downward-propagating ELF/VLF wave reflects at the ground, travels back to the ionosphere, and then reflects again in the ionosphere to make a second downward-propagating wave. As a result, there is strong constructive interference when the distance between the ground and the reflection altitude is a multiple of half an ELF/VLF wavelength. This occurs at harmonics of approximately 2 kHz. To put this in quantitative terms, if $\mathbf{H}_0$ is the primary wave, R_{EI} is the product of the earth and ionosphere reflection coefficients, z_R is the reflection height, and we ignore subsequent reflections ($R_{EI} \ll 1$), then the downward-going wave is of the form

$$\mathbf{H} = (1 + R_{EI}e^{2ikz_R})\mathbf{H}_0. \tag{4.24}$$

A nonzero reflection coefficient R_{EI} causes periodic bumps in the measured magnitude and phase spectrum of $\mathbf{H}$. The peak-to-peak size of the magnitude and phase bumps is given by:

$$\frac{\Delta A}{A} = 2R_{EI} \tag{4.25}$$

$$\Delta\psi = 2\arcsin R_{EI}. \tag{4.26}$$

With the above influences in mind, we plot the magnitude of the polarization ellipse A and its phase ψ. These are shown in Figures 4-15 and 4-16, respectively.

The amplitude spectra have been normalized to the level of the 7285 or 7290 Hz calibration pulse transmitted every 10 seconds (see Section 4.2). The normalizing removes the dependence of the amplitude on the variation of $\mathbf{E}^0$ during the 20 or 30 minutes required for the modulation frequency sweep. The amplitude spectra are on the whole rather flat except for the resonance peaks at multiples of 2 kHz due to the cavity resonance, and a small rolloff at each end. The peak-to-peak amplitude of the bumps is about 2 or 3 times smaller than the total amplitude, which puts the reflection coefficient R_{EI} in the range 1/6 to 1/4. The flat spectrum is interpreted as due to the cancellation of the factor k in Equation (4.17) with the factor of $1/\omega_m$ in Equation (3.13). The rolloff below 2 kHz is interpreted as due to the transition from the fast modulation limit [Eq. (3.13)] to the slow modulation limit [Eq. (3.3)]. The rolloff above 6 kHz is interpreted as due to the phase mixing of radiation from different altitudes.

The phase spectra are dominated by the linear phase term e^{ikz} due to the separation of the radiation source and the observer. More accurately speaking, since we are taking the phase of the modulation at the HAARP transmitter to be the phase reference, the phase results from the trip up and down to the heated layer, and is therefore of the form e^{2ikz}. The slope of the phase:

$$\frac{\partial}{\partial \omega}(2kz) = \frac{2z}{c} \tag{4.27}$$

gives the "apparent altitude" z of the source region, which is the assumption that the propagation in both directions occurs at c. z is computed by fitting the phase data between 2 and 8 kHz to a line of the form $kz + \psi_0$ via linear least-squares. This calculation results in an estimate for z (the slope). The result for z is inset on the plots in Figure 4-16.

Some authors [*Stubbe et al.*, 1981] choose to differentiate the phase data directly and get an apparent altitude as a function of modulation frequency. However, dif-
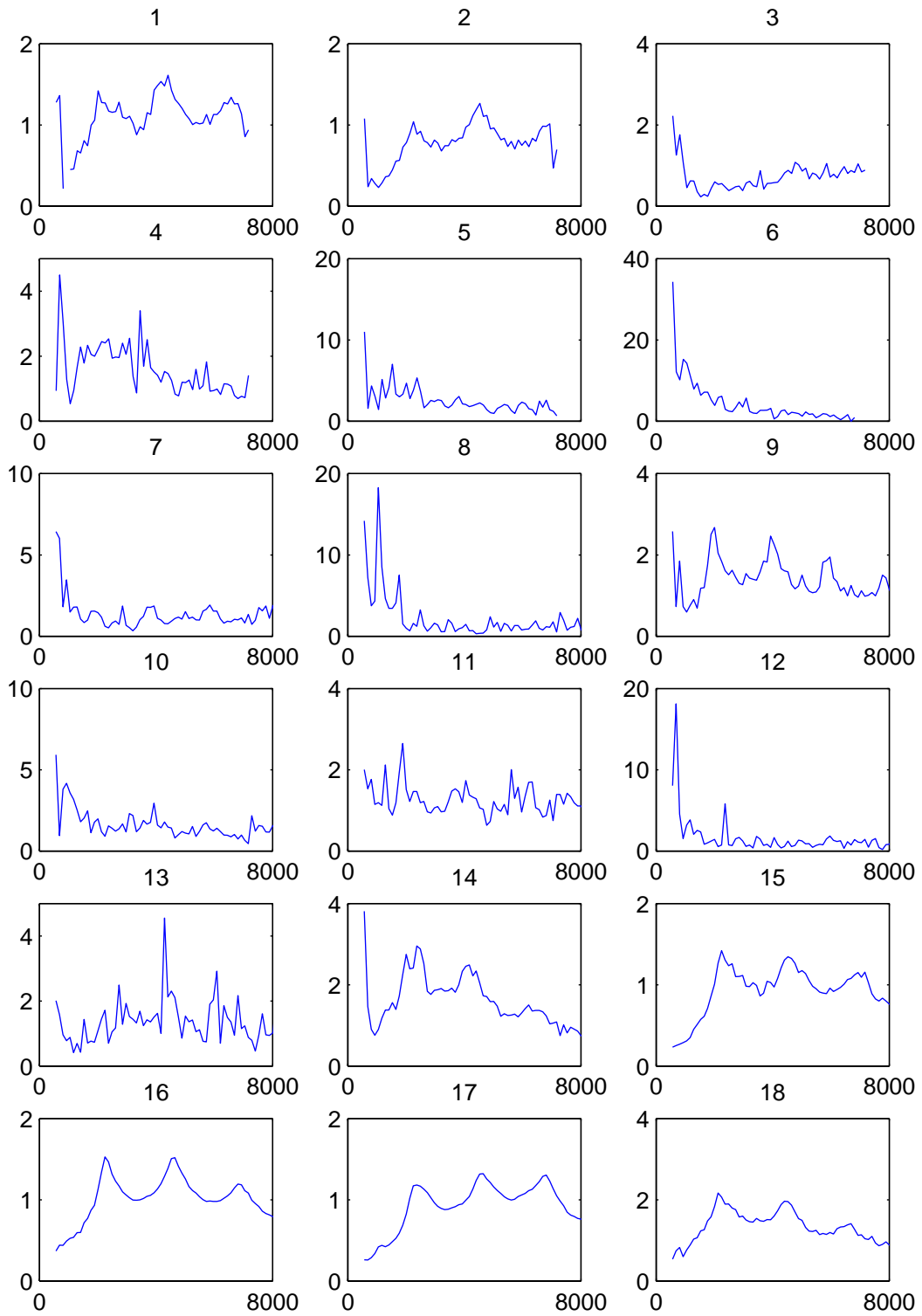
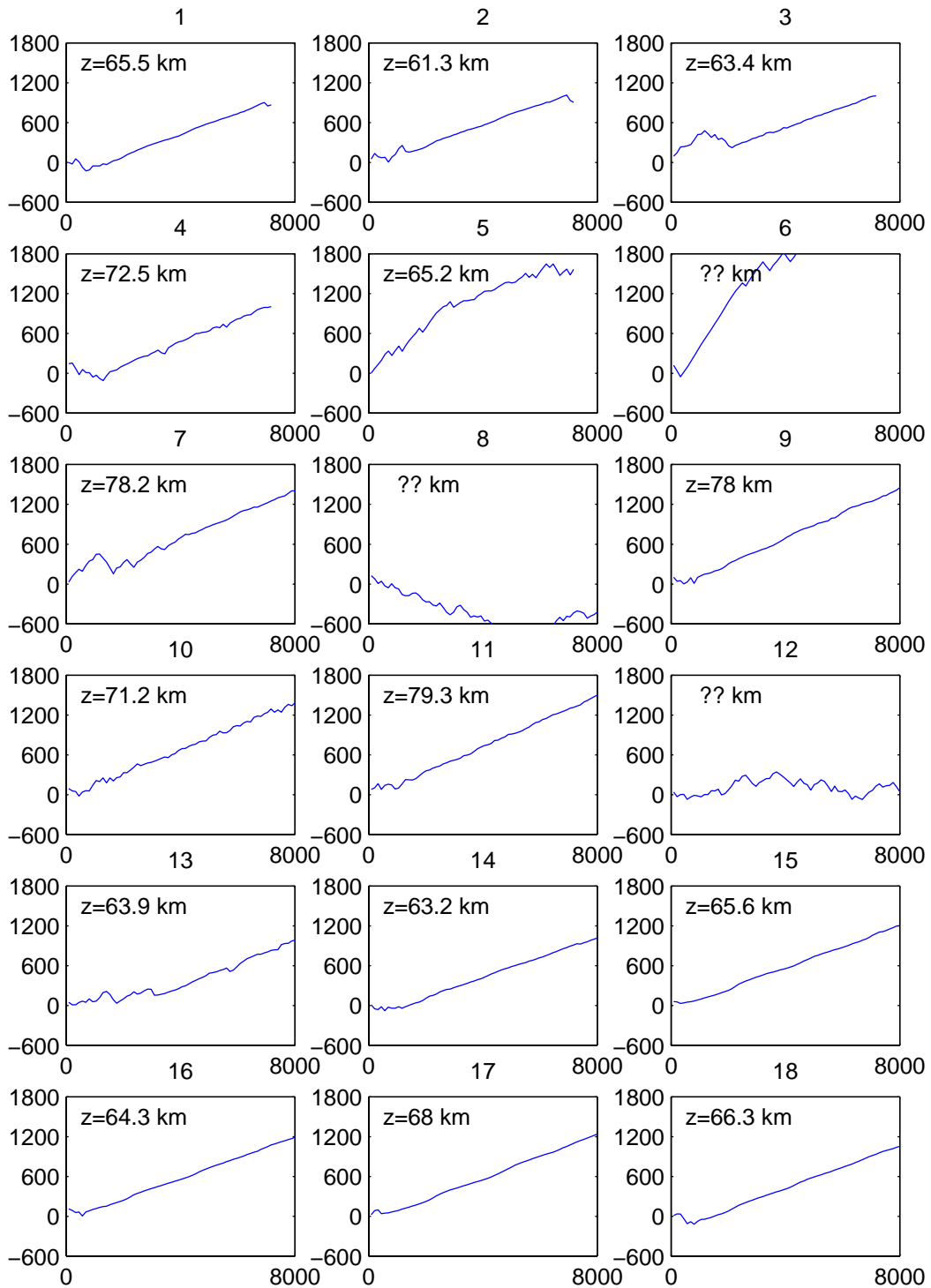Figure 4-15: Normalized magnitude response versus modulation frequency (Hz).

Figure 4-16: Phase ψ (degrees) versus modulation frequency (Hz). Apparent radiation source altitude indicated.

130

ferentiating data is problematic due to noise, and even in circumstances where the derivative can be made, the presence of undulations in the phase $\Delta\psi$ due to the cavity resonance render the results difficult to interpret. Figure 4-17 shows the outcome of such a differentiation. Indeed the results are quite variable, although in a few cases the altitude is readable, with peaks at 2 kHz harmonics due to $\Delta\psi$.

A futher point here is that the differentiation disregards the constant phase ψ_0 across the band. To illustrate this point, in Figure 4-18 we subtract the fitted linear phase kz from the data and plot the remaining phase. Almost half of the valid data sets show a significant component of constant negative phase, in some cases approaching -90 degrees. This observation has not been previously reported in the literature. The interpretation of this result is that there can exist a current profile with a significant doublet component.

A more exact statement concerning the implications of this phase can be made by actually inverse-transforming the experimental data to obtain the current profile, following Equation (4.17). Unfortunately, these particular experiments were not actually designed with this operation in mind, and thus there is a severe limitation to the spatial resolution, since the frequency spectra only extend out to 8 kHz (corresponding to $k = 1.6 \times 10^{-4}$ m^{-1}). Generally speaking, features of size smaller than half a wavelength at 8 kHz cannot be resolved. In other words, a structure in the current profile smaller than 20 km in size cannot be determined with the inverse Fourier transform. However, we can take advantage of the very good experimental conditions during the overhead current jet (shown in Figure 4-10) when ELF/VLF radiation was measurable at higher harmonics of the modulation frequency.

During these conditions, one can exploit the fact that the square wave heater modulation function $p(t)$ contains third and fifth harmonics at power levels 1/3 and 1/5 of the fundamental, respectively. In the fast modulation limit, the heating process is essentially linear [Eq. (3.13)], so radiation at the third and fifth harmonic of the modulation frequency is produced independently of the fundamental. Therefore one has essentially additional experiments at all the harmonics of the modulation frequency, and measurements of the harmonic radiation can be interpreted in exactly
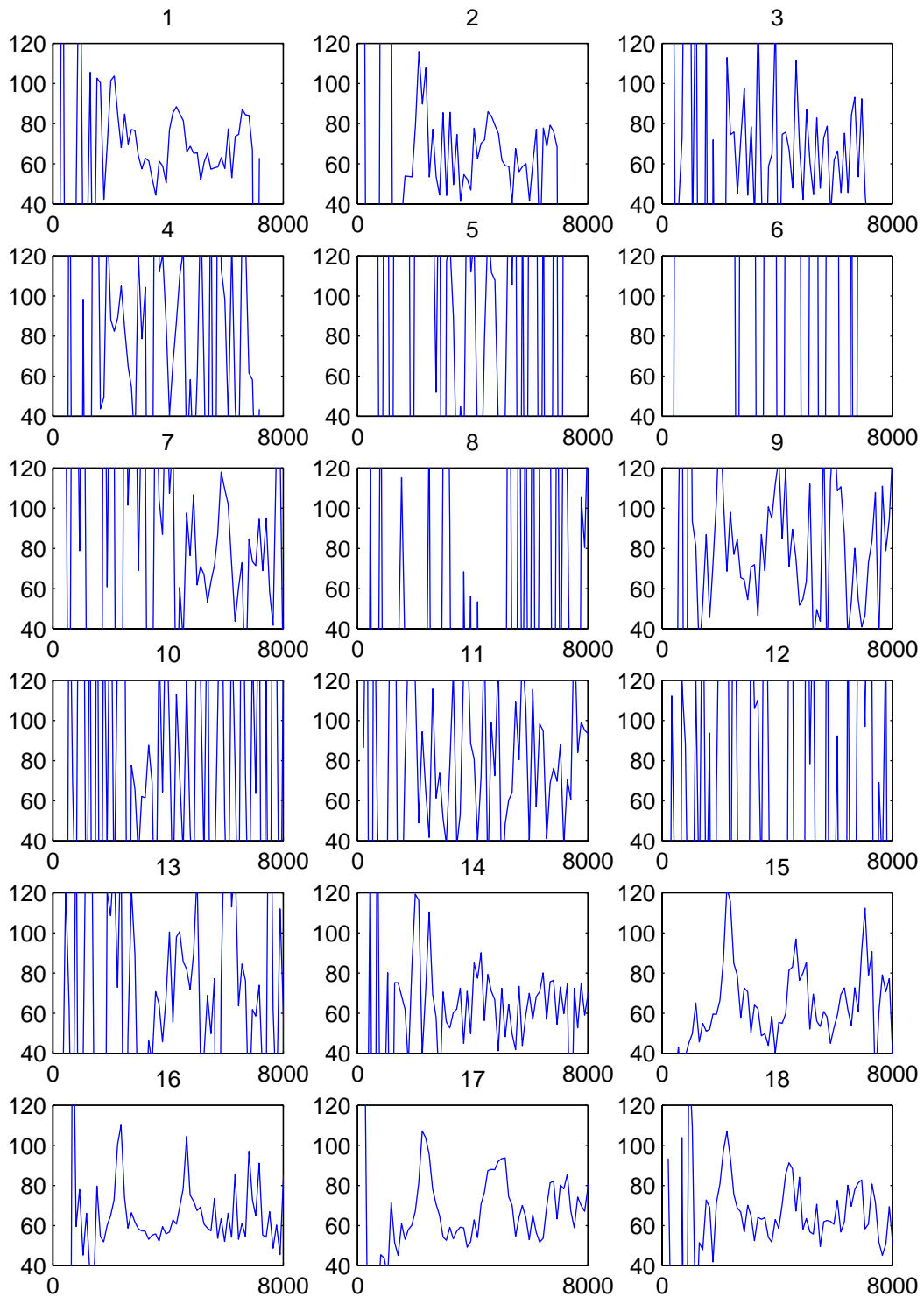
Figure 4-17: Apparent altitude (km) of radiation as a function of modulation frequency (Hz), obtained by differentiation of the phase.
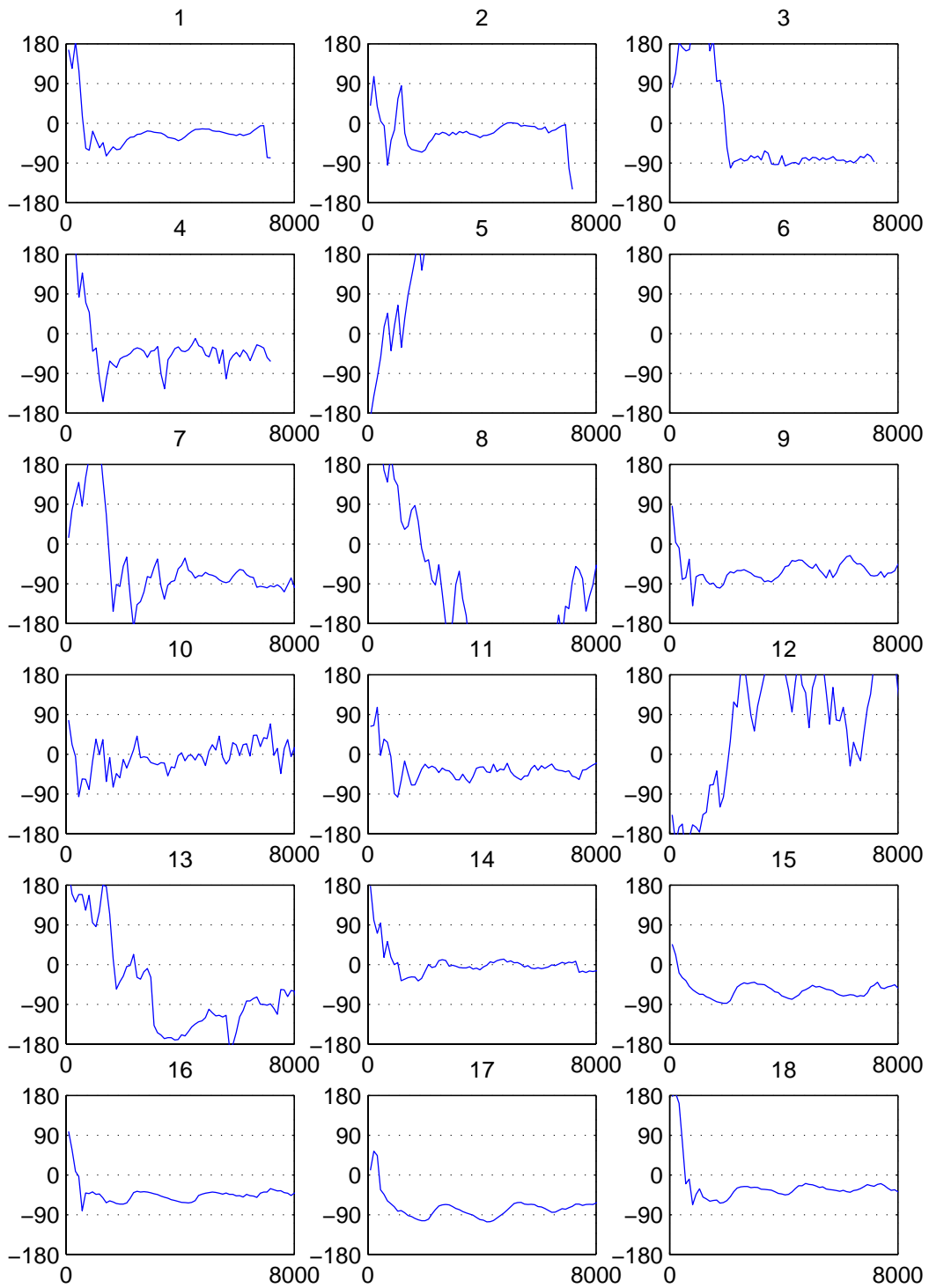
Figure 4-18: Departure from linear phase (degrees) versus modulation frequency (Hz).
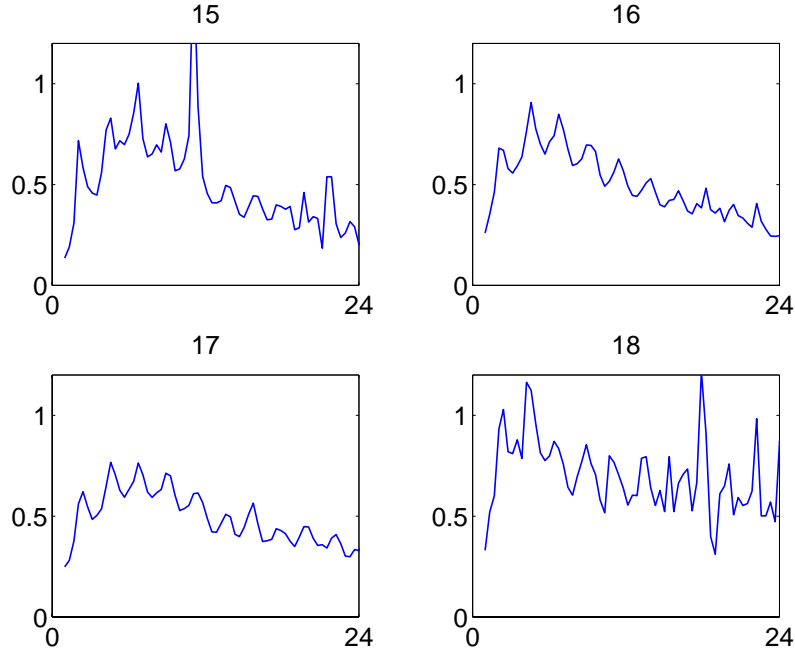
133

Figure 4-19: Normalized magnitude of third harmonic magnetic field versus modulation frequency (kHz).

the same way as the fundamental. While third and fifth harmonic radiation would extend the spectrum to 24 and 40 kHz, respectively, the downside is that the power levels are low, and the radiation is observable in but a few of the data sets.

We consider data sets (15,16,17,18). The magnitude A and phase ψ of these data sets at the third and fifth harmonics are given in Figures 4-19 through 4-22.

The first, third, and fifth harmonics of the east-west and north-south magnetic field components of data sets (15,16,17,18) are inverse Fourier-transformed to produce the experimentally determined profiles given in Figures 4-23 through 4-25. The left column in each is the x-directed current, and the right column is the y-directed current. In performing the inverse transforms on the experimental data, the linear phase factor is lumped into $\tilde{J}_{cx}$ and $\tilde{J}_{cy}$ [see Equation (4.17)]. This inclusion causes the deduced current structures to appear at the "apparent altitude" when the inverse transform is carried out.

The deduced structure is double-layered, with two oppositely directed current layers approximately 10 km apart. Both the J_x and J_y components show almost
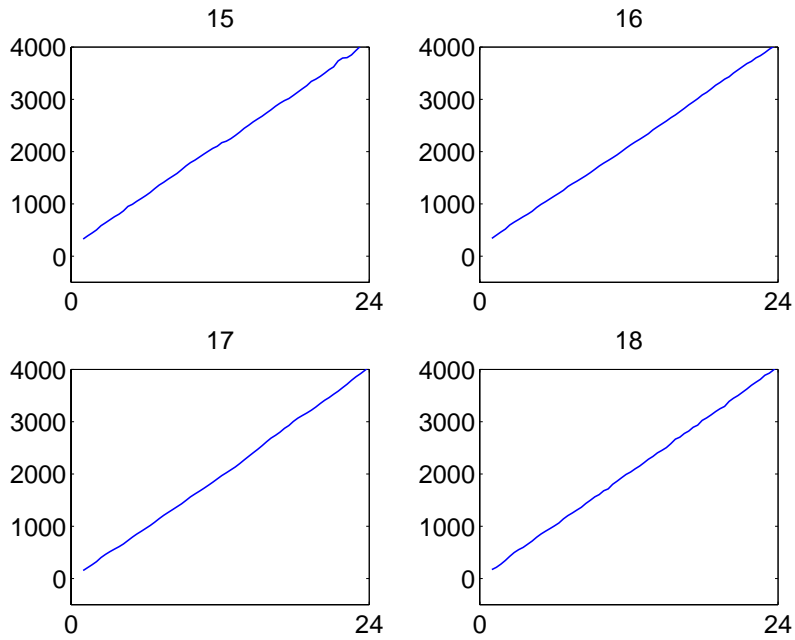
134

Figure 4-20: Phase of third harmonic magnetic field (degrees) versus modulation frequency (kHz).
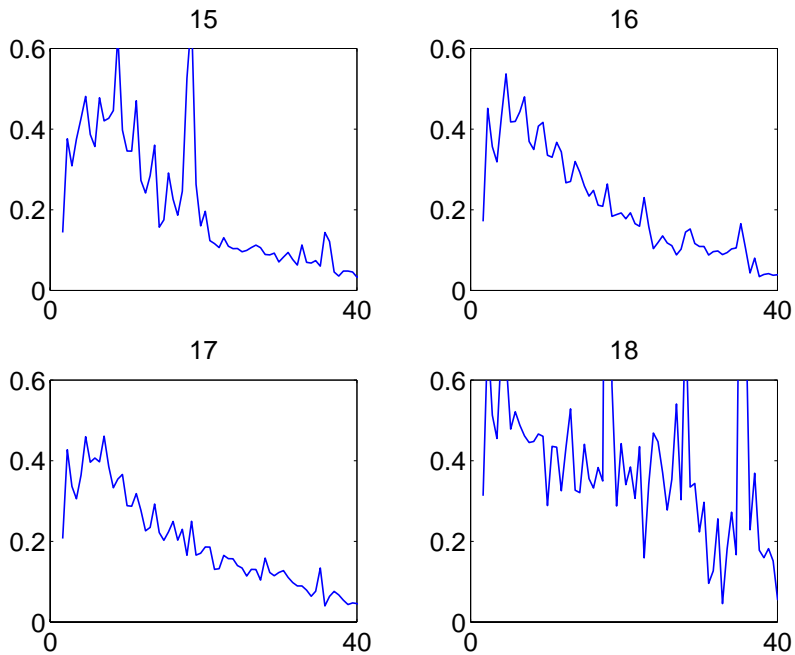

Figure 4-21: Normalized magnitude of fifth harmonic magnetic field versus modulation frequency (kHz).
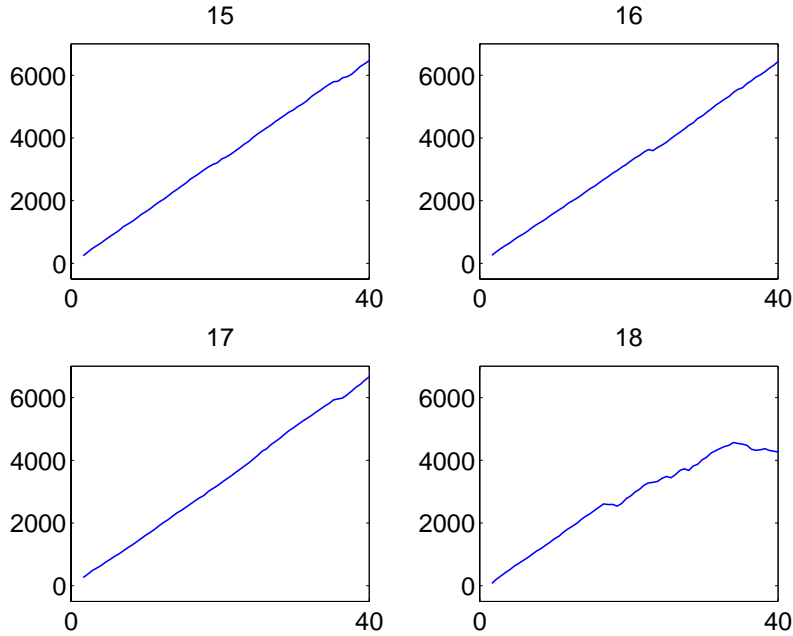
Figure 4-22: Phase of fifth harmonic magnetic field (degrees) versus modulation frequency (kHz).

identical structure, with the lower layer considerably wider than the upper layer. The currents are directed such that both J_x and J_y are negative in the lower layer and both are positive in the upper layer. This implies that the lower currents are in the physical southwest direction, and the upper currents are in the northeast direction.

A different view of the current profiles is given in Figure 4-26. Here the two components of the deduced current are plotted together as a function of altitude. Note the rapid counterclockwise reversal of the current direction near the altitude of 68 km. The direction reverses within about 3 or 4 km.

We are now in a position to compare the observations with the theory. The structure of the current in the vertical dimension is obviously not predicted by the two-dimensional theory of *Stubbe and Kopka* [1977] (Section 3.3.2) and thus we can immediately regard this theory as inadequate. Next we ask, what can be said of the theory of *Stubbe et al.* [1982] (Section 3.3.1)? To compare the theory to the experiment, we need to rotate the theory counterclockwise by about 90 degrees to account for the fact that $\mathbf{E}^0$ is presumed physically northward during the time of
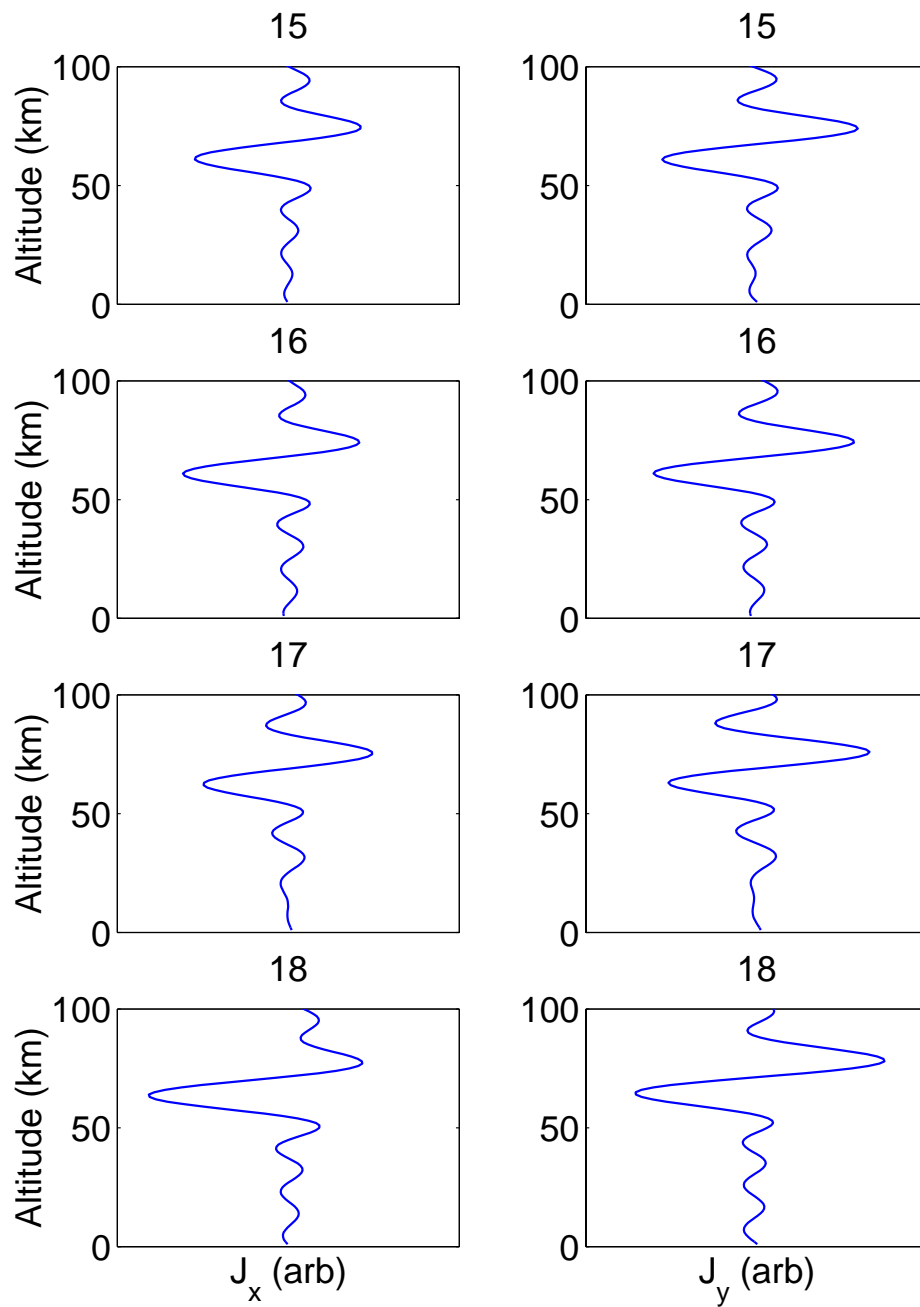
Figure 4-23: Current profile deduced via inverse Fourier transform of first harmonic data (note poor resolution compared to 3rd, 5th harmonics).
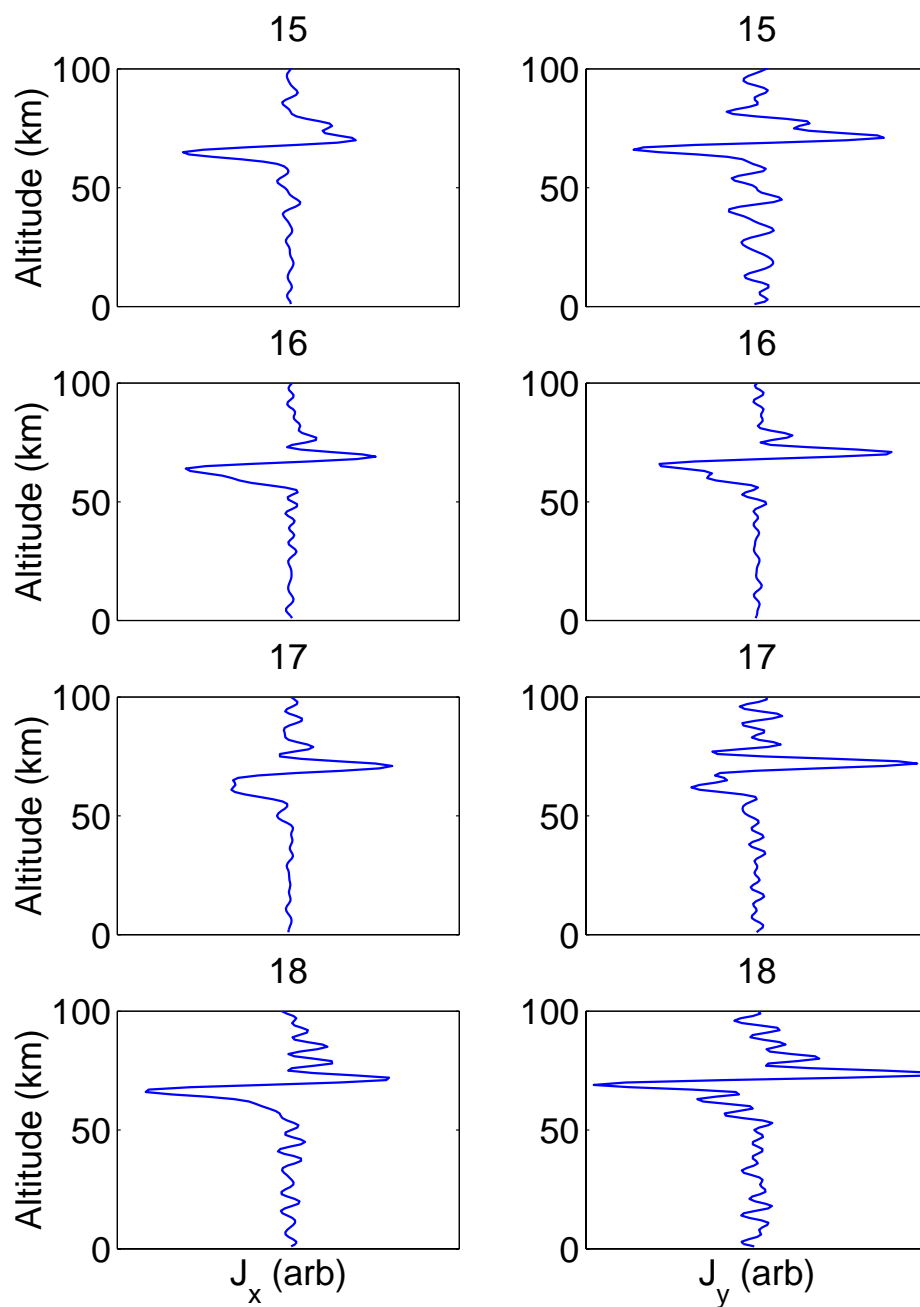
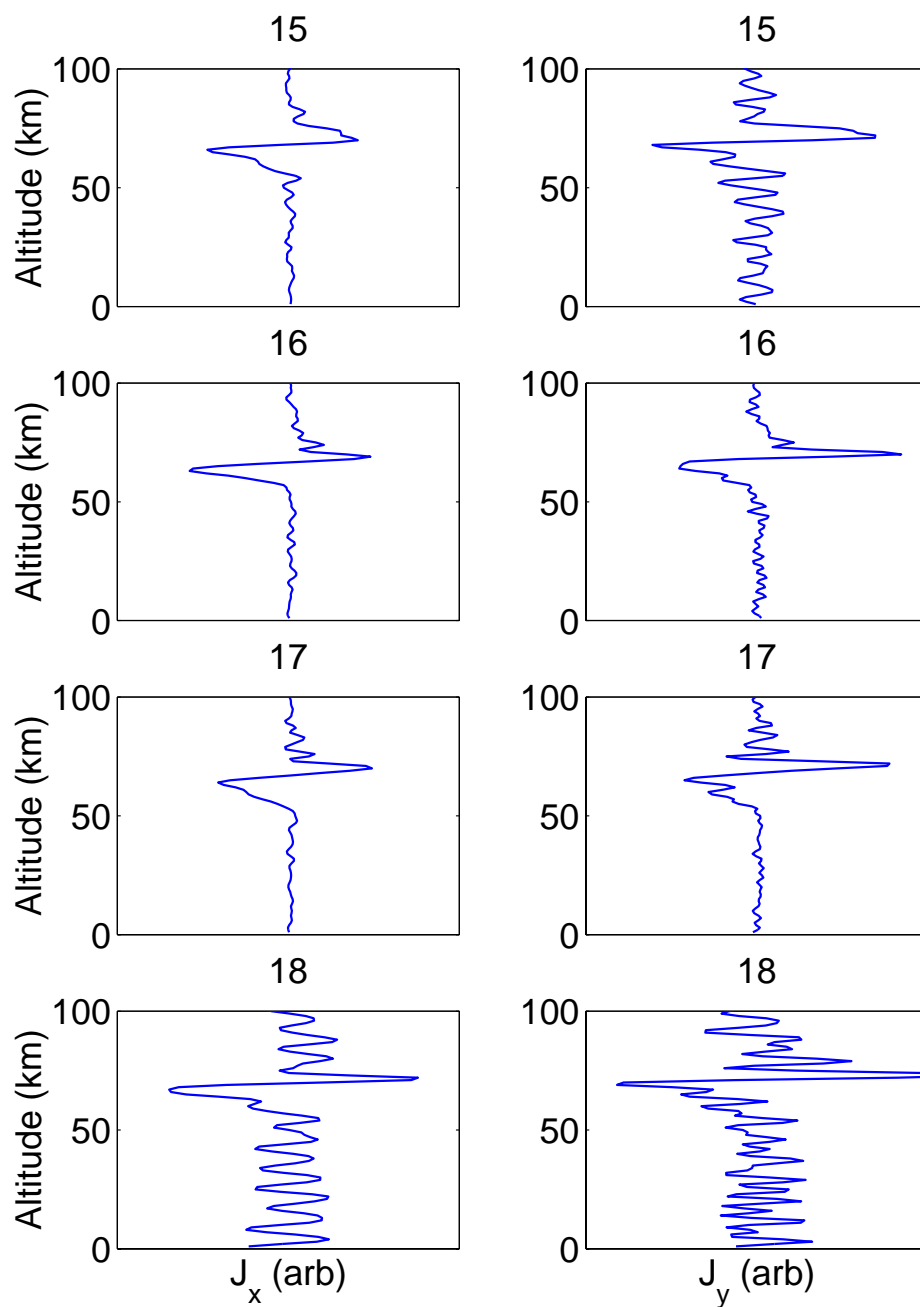Figure 4-24: Current profile deduced via inverse Fourier transform of third harmonic data.

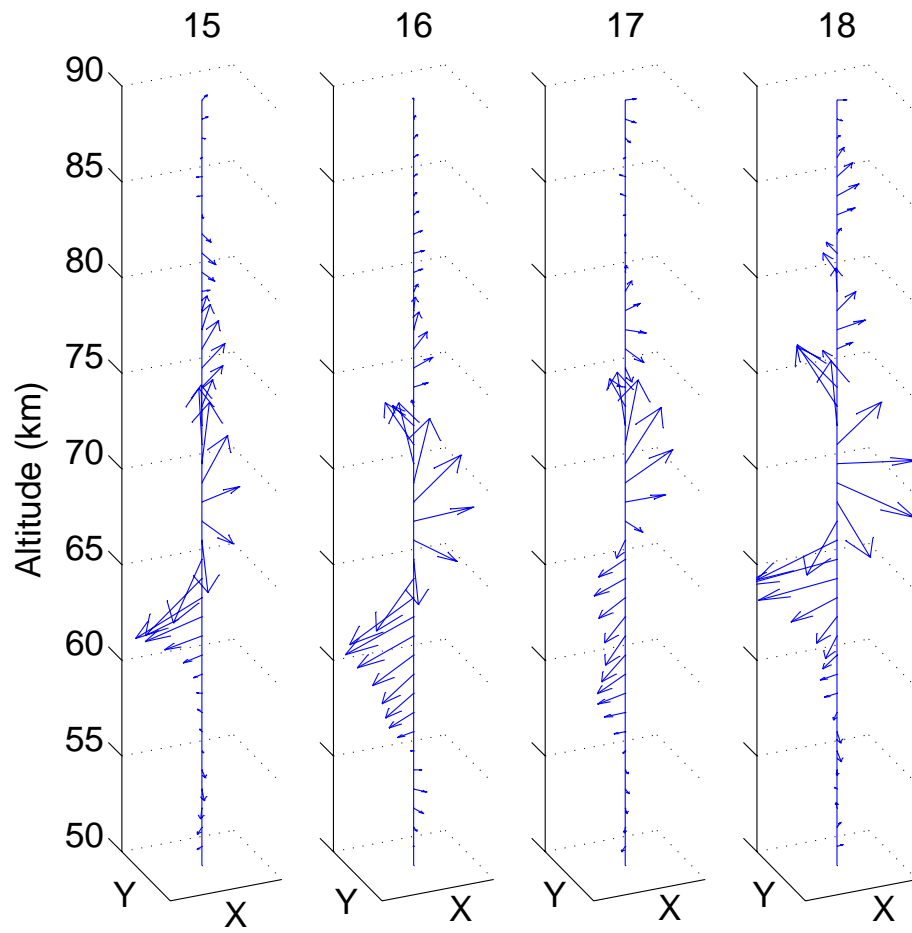Figure 4-25: Current profile deduced via inverse Fourier transform of fifth harmonic data.

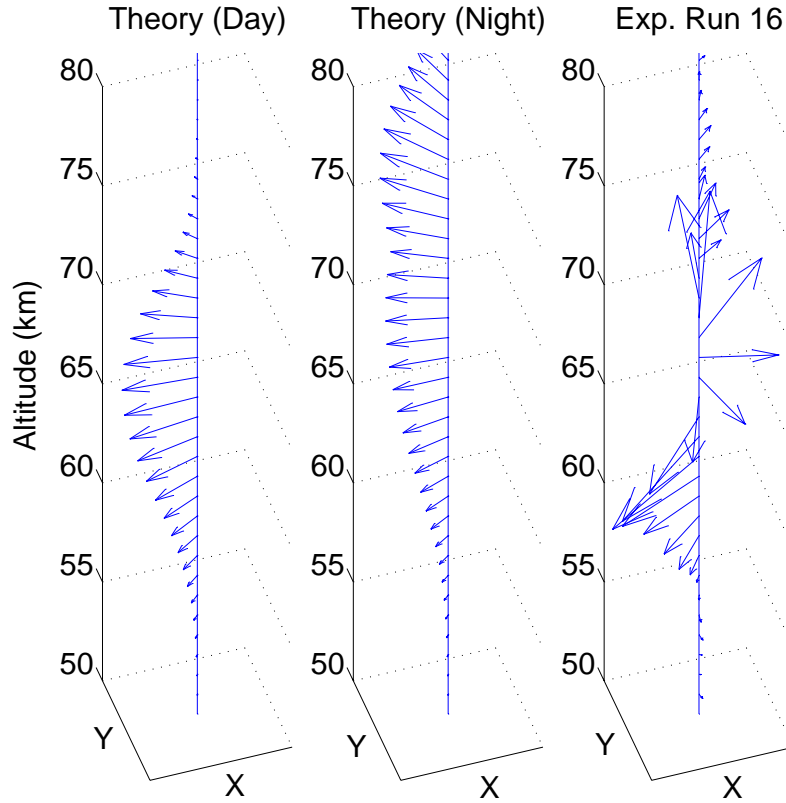Figure 4-26: Current vector versus altitude (from third harmonic data).

Figure 4-27: *Stubbe et al.* [1982] theory compared to experimentally determined current vector.

these measurements (see Section 4.3.2) rather than eastward, as in the theory. In Figure 4-27 we compare the theoretical profiles of Section 3.3.1 with data set 16. The theory of Section 3.3.1 basically says that the x-component of the current will be proportional to the Pedersen conductivity change and the y-component of the current will be proportional to the negative of the Hall conductivity change. Since we are interested in fast modulation, we use the profiles represented in Figure 3-14.

Neither the daytime nor nighttime fluctuating conductivity profiles give rise to a current profile which has any resemblance to the measured profile. In the theory, the current rotates very slowly in the x-y plane compared to the sudden direction change of the experimental data. Thus the theory of *Stubbe et al.* [1982] (Section 3.3.1) cannot explain the experimental results. Extensions to the theory, as suggested in Sections 3.3.3 and 3.3.4, are therefore required.
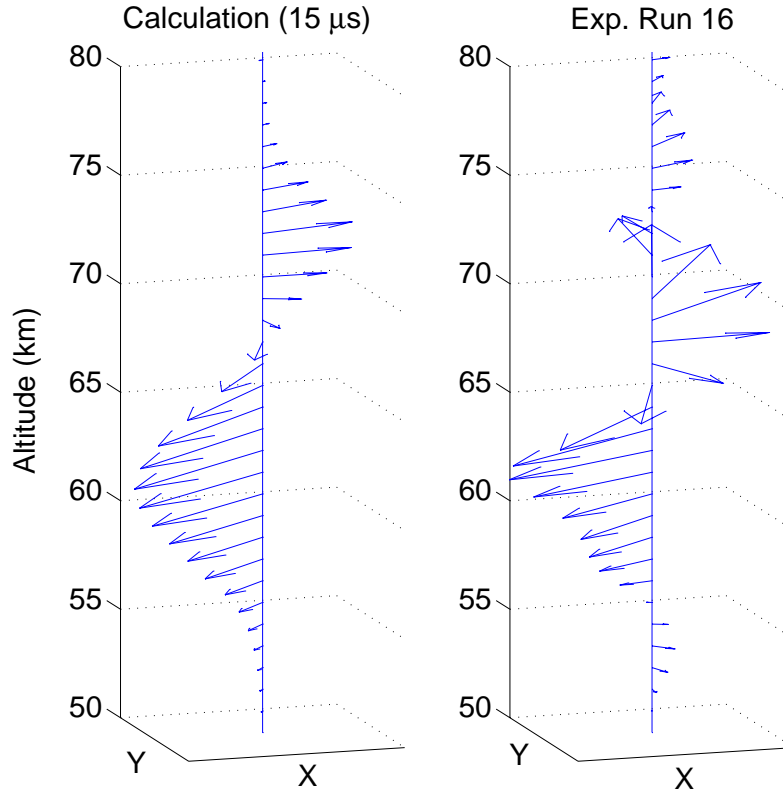
Figure 4-28: Numerical solution compared to experimentally determined current vector.

While the third level theory (Section 3.3.3) predicts vertical structure, it does so only crudely as it is height-integrated in two layers. A more satisfactory comparison can be made between the data and the numerical calculations (Section 3.3.4). We note that the experiment conditions are nominally nighttime, while the simulation is nominally daytime. Nevertheless, the strong riometer absorption, low apparent altitude (64–68 km), and the appearance of visible aurora during experimental runs (15,16,17,18) all suggest an unusually strong density profile more akin to the daytime conditions. This issue notwithstanding, a comparison between the calculations and the data is done in Figure 4-28. Here we see that there is now reasonable qualitative agreement between theory and experiment with regards to the altitude of the current reversal, the azimuthal direction of the current, the spatial rate at which the current vector rotates, and the sense of the current rotation. The chosen 15 μs

142

snapshot captures the current structure at a time scale which is representative of the midpoint of the frequency range (10 kHz) over which the third harmonic of the modulation frequency is swept. Snapshots at longer times will produce even sharper current rotations near 68 km. The issue is not that important, however, as the inverse transform-deduced current profile is essentially a frequency integrated quantity and thus represents the average character of the current structure over the entire frequency band. Frequency integrated theory will not be pursued here as the vertical loop-like average character has already been clearly established in Section 3.3.4.

Chapter 5

Conclusion

A comprehensive description of lower ionospheric heating in the presence of solar dynamo electric fields has been presented in Chapters 2 and 3. Theories concerning the generation of ELF/VLF antenna current by modulated heating (Sections 3.3.1–3.3.4) have been examined. It has been shown that two often-used theories [*Stubbe et al.*, 1982, *Stubbe and Kopka*, 1977] (Sections 3.3.1, 3.3.2) cannot explain the two-layer vertical structure of antenna current deduced from the author's HAARP experimental results (Chapter 4). The author has proposed analytic and numerical extensions to the theories (Sections 3.3.3, 3.3.4) which can explain the qualitative features of the deduced vertical current structure. Specifically, the extensions predict the observed rapid directional change with altitude of the current associated with the deduced vertical loop structure.

The major limitation of this work is the need for good natural conditions (strong electrojet current and a low-altitude ionospheric profile) in order to produce the wide frequency spectrum necessary to resolve the current structure. Thus the results in this work are for unusually good conditions, and not for the more mundane conditions to be encountered in the everyday circumstances of ELF/VLF communication work. However, the qualitative predictions of the theory hold for small, height-integrated perturbations to the background conductivity, and thus should hold for weak electrojet current and weak ionospheric density conditions.

A good direction for future work would be to run the simulations under a large

variety of natural conditions to verify the assertion of the previous sentence. However, to get experimental results under ordinary conditions, a different experimental setup would be required. One possibility is to deploy a number of VLF receivers underneath the heated volume and get phase information by measuring the radiation at different points in space rather than at different radiation frequencies. The current structure could then be deduced at a single modulation frequency rather than frequency integrated as was done in this thesis. Such a multiple-receiver system is currently being deployed at HAARP, and in principle will allow the real-time measurement of three-dimensional current structure during even modest ionospheric conditions.

Appendix A

Magnetic Induction Analytic Model

Here we consider the problem of the vertical loop height in detail (refer to Section 3.3.3). To facilitate a simple solution to this problem, we employ rectangular geometry as per the following assumptions (refer to Figure A-1):

1. The heated volume is a box with linear dimensions l, w and h.

2. The conductivity modulation causes the plasma in the box to polarize at an angle $\alpha = -\arctan(\Delta\Sigma_H/\Delta\Sigma_P)$ with respect to $\mathbf{E}^0$.

3. The box is rotated at an angle α with respect to $\mathbf{E}^0$, such that the conductivity modulation causes a uniform layer of polarization charge to appear at each end of the box.

4. The potential difference across the box is U_0.

In the context of the above geometry, we now determine the potential variation $U(z)$ in the region above the box. If we consider the modulation current leaving one end of the box to be $I_0 = 2wD\Sigma_P''$, then the variation in vertical current $I(z)$ will depend on the vertical distribution of return current above the box. The change in vertical
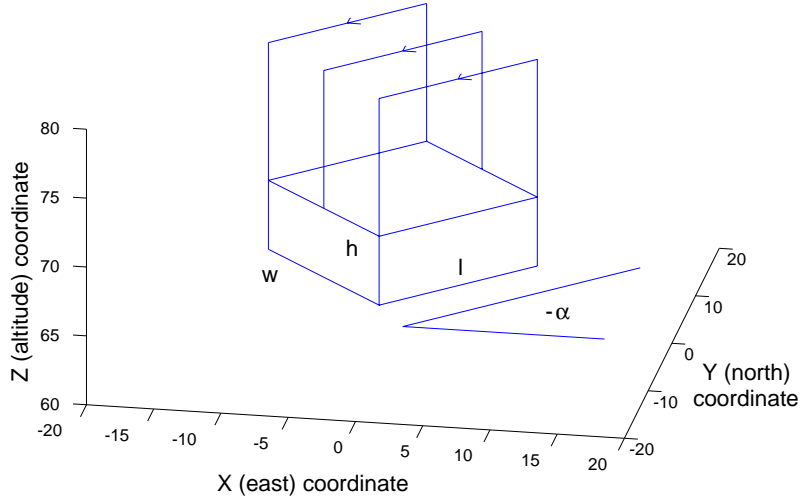
Figure A-1: Box geometry.

current with altitude is given by

$$\frac{\partial I}{\partial z} = -G_z U, \tag{A.1}$$

where G_z is the horizontal conductance per unit height. Similarly, the loop self-inductance will cause a back-EMF which will result in U decaying with height. The variation is given by

$$\frac{\partial U}{\partial z} = -L_z \frac{\partial I}{\partial t}, \tag{A.2}$$

where L_z is the self-inductance of the current loop per unit height. Combining these equations gives us the relation

$$\frac{\partial^2 U}{\partial z^2} - L_z G_z \frac{\partial U}{\partial z} = 0, \tag{A.3}$$

with $U = U_0$ at the top of the heated volume, taken as $z = 0$. If we ignore fringing currents, then the crossfield conductance per unit height is

$$G_z = \frac{\sigma_p w}{l}. \tag{A.4}$$

148

and the loop inductance per unit height is

$$L_z = \frac{\mu_0 l}{w}. \tag{A.5}$$

Taking $\sigma_P = \sigma_{P0} e^{z/d}$ (see Figure 3-11), where d is the scale height of the conductivity, and letting $\frac{\partial}{\partial t} \rightarrow -i\omega_m$, we have

$$\frac{\partial^2 U}{\partial z^2} + i\omega_m \mu_0 \sigma_{P0} e^{z/d} U = 0. \tag{A.6}$$

The general solution to this homogeneous differential equation is

$$U = C_1 J_0 \left(2d\, e^{z/2d} \sqrt{i\omega_m \mu_0 \sigma_{P0}} \right) + C_2 Y_0 \left(2d\, e^{z/2d} \sqrt{i\omega_m \mu_0 \sigma_{P0}} \right), \tag{A.7}$$

where J_0 and Y_0 are zero-order Bessel functions of the first and second kind, and C_1 and C_2 are arbitrary constants used to match the boundary conditions. We need U to remain bounded as $z \rightarrow \infty$. Considering the limiting forms of the Bessel functions for large arguments:

$$J_0(z) = \sqrt{\frac{2}{\pi z}} \cos \left(z - \frac{\pi}{4} \right) \tag{A.8}$$

$$Y_0(z) = \sqrt{\frac{2}{\pi z}} \sin \left(z - \frac{\pi}{4} \right), \tag{A.9}$$

we need $C_1 = -iC_2$ for U to remain bounded. Thus U can be written as a Hankel function of the first kind with a single constant:

$$U = C H_0^{(1)} \left(2d\, e^{z/2d} \sqrt{i\omega_m \mu_0 \sigma_{P0}} \right). \tag{A.10}$$

To find C, we apply current conservation at the boundary of the heated volume. The current flowing out of the heated volume is just $I_0 = I(z = 0)$. From Equation (A.2) we have

$$I_0 = \frac{1}{i\omega_m L_z} \frac{\partial U}{\partial z} \bigg|_{z=0} \tag{A.11}$$

149

$$= C\sqrt{\frac{i^3\sigma_{P0}}{\omega_m\mu_0}}H_1^{(1)}\left(2d\sqrt{i\omega_m\mu_0\sigma_{P0}}\right). \tag{A.12}$$

In most cases, the Hankel function argument is less than one and can be expanded as $H_1^{(1)}(z) \approx -2i/\pi z$:

$$I_0 = \frac{C}{\pi\omega_m\mu_0 d}. \tag{A.13}$$

This current has to be equal to the modulation current impinging on the side of the rectangular heated volume:

$$I_0 = 2wD\Sigma_P''. \tag{A.14}$$

Comparing this expression with Equation (A.13) gives us

$$C = 2\pi\omega_m\mu_0 wdD\Sigma_P''. \tag{A.15}$$

Therefore the return modulation current per unit height I_z above the heated volume is given by

$$I_z = UG_z \tag{A.16}$$

$$= \frac{\sigma_{P0}w}{l}e^{z/d}CH_0^{(1)}\left(2d\,e^{z/2d}\sqrt{i\omega_m\mu_0\sigma_{P0}}\right). \tag{A.17}$$

This is a complex amplitude which needs to be separated into real and imaginary parts. This decomposition is facilitated by invoking the Kelvin functions

$$\ker x + i\ker x = -\frac{i\pi}{2}H_0^{(1)}\left(x\sqrt{i}\right). \tag{A.18}$$

These functions are shown in Figure A-2. We can express the time-varying current per unit height above the heated volume in terms of these functions as follows

$$I_z = \frac{2\sigma_{P0}wCe^{z/d}}{\pi l}\left[\ker\left(2de^{z/2d}\sqrt{\omega_m\mu_0\sigma_{P0}}\right)\sin\omega_m t \right. \tag{A.19}$$

$$\left. -\kei\left(2de^{z/2d}\sqrt{\omega_m\mu_0\sigma_{P0}}\right)\cos\omega_m t\right]. \tag{A.20}$$
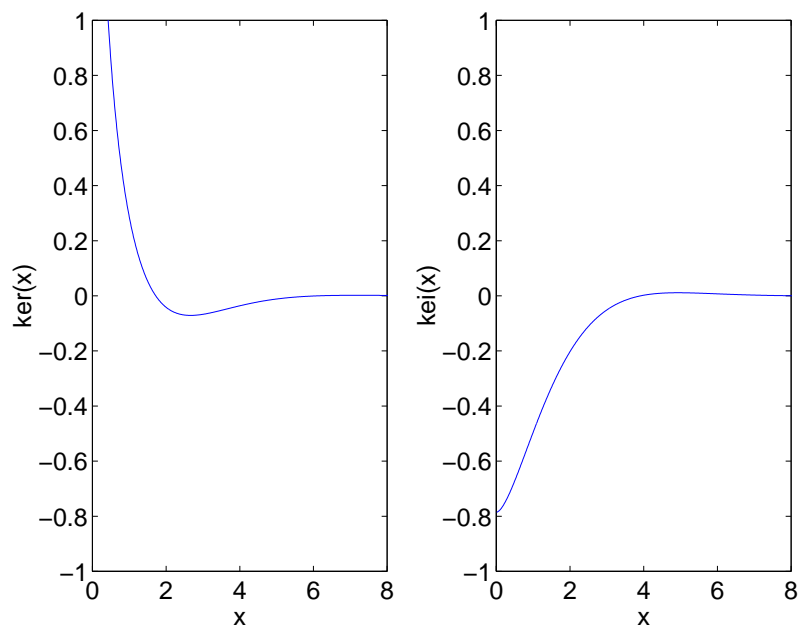
Figure A-2: Kelvin functions.

Bibliography

[1] Allis, W. P., Motions of ions and electrons, in *Encyclopedia of Physics*, vol. 21, edited by S. Flugge, 383, Springer, Berlin, 1956.

[2] Barr, R., and P. Stubbe, ELF and VLF radiation from the "polar electrojet antenna", *Radio Sci., 19*, 1111, 1984.

[3] Barr, R., P. Stubbe, M. T. Rietveld, and H. Kopka, ELF and VLF signals radiated by the "polar electrojet antenna": experimental results, *J. Geophys. Res., 91*, 4451, 1986.

[4] Crompton, R. W., and D. J. Sutton, Experimental investigation of the diffusion of slow electrons in nitrogen and hydrogen, *Proc. Roy. Soc. London, 215A*, 467, 1952.

[5] Engelhardt, A. G., A. V. Phelps, and C. G. Risk, Determination of momentum transfer and inelastic collision cross sections for electrons in nitrogen using transport coefficients, *Phys. Rev., 135A*, 1566, 1964.

[6] Frost, L. S., and A. V. Phelps, Rotational excitation and momentum transfer cross sections for electrons in H_2 and N_2 from transport coefficients, *Phys. Rev., 127*, 1621, 1962.

[7] Gerjuoy, E., and S. Stein, Rotational excitation by slow electrons, *Phys. Rev., 97*, 1671, 1955.

[8] Getmantsev, G. G., A. V. Gul'el'mi, B. I. Klain., D. S. Kotik, S. M. Krylov, N. A. Mityakov, V. O. Rapoport, Yu. A. Sazonov, V. Yu. Trakhtengerts, and V. Ya. Eidman, *Pis'ma Zh. Eksp. Teor. Fiz., 20*, 229, 1974.

[9] Gurevich, A. V., Nonlinear phenomena in the ionosphere, Springer-Verlag, New York, 1978.

[10] Hake, R. D., and A. V. Phelps, Momentum-transfer and inelastic collision cross sections for electrons in O_2, CO, and CO_2, *Phys. Rev., 158*, 70, 1967.

[11] Herzberg, G., Spectra of diatomic molecules, D. van Nostrand Company, Inc., Princeton, New Jersey, 1950.

[12] James, H. G., R. L. Dowden, M. T. Rietveld, P. Stubbe, and H. Kopka, Simultaneous observations of ELF waves from an artificially modulated auroral electrojet in space and on the ground, *J. Geophys. Res., 89*, 1655, 1984.

[13] Jespersen, M., A. Haug, and B. Landmark, Electron densities and collisional frequency observations in the arctic D-region, in *Electron Density Profiles in the Ionosphere and Exosphere*, edited by J. Frighagen, 27, North-Holland, New York, 1966.

[14] Johnson, F. S., Satellite Environment Handbook, Stanford Univ. Press, Stanford, 1961.

[15] Kasputin, N. N., R. A. Pertsovskii, A. N. Vasil'ev, O. M. Raspopov, D. A. Ul'yanchenko, and A. A. Arykov, *Pis'ma Zh. Eksp. Teor. Fiz., 25*, 248, 1977.

[16] Kelley, M. C., The earth's ionosphere, Academic Press, Inc., San Diego, 1989.

[17] Kotik, D. S., and M. A. Itkina, On the physical limit of the power of heating facilities, *J. Atmos. Terr. Phys., 60*, 1247, 1998.

[18] Kuo, S. P., S. H. Lee, and P. Kossey, Major enchancement of extra-low-frequency radiation by increasing the high-frequency heating wave power in electrojet modulation, *Phys. Plasmas, 9*, 315, 2002.

[19] Milikh, G. M., M. J. Freeman, and L. M. Duncan, First estimates of HF-induced modifications of the D region by the HF Active Auroral Research Program facility, *Radio Sci., 29*, 1355, 1994.

[20] Milikh, G. M., K. Papadopoulos, M. McCarrick, and J. Preston, ELF emission generated by the HAARP HF-heater using varying frequency and polarization, *Radiophysics and Quantum Electronics, 42*, 639, 1999.

[21] Papadopoulos, K., C. L. Chang, P. Vitello, and A. Drobot, On the efficiency of ionospheric ELF generation, *Radio Sci., 25*, 1311, 1990.

[22] Papadopoulos, K., H.-B. Zhou, and C.-L. Chang, Cerenkov excitation of whistler/ helicon waves by ionospheric HF heating, *Geophys. Res. Lett., 21*, 1767, 1994.

[23] Rietveld, M. T., H. Kopka, and P. Stubbe, D-region characteristics deduced from pulsed ionospheric heating under auroral electrojet conditions, *J. Atmos. Terr. Phys., 48*, 311, 1986.

[24] Rietveld, M. T., H.-P. Mauelshagen, P. Stubbe, H. Kopka, and E. Nielsen, The characteristics of ionospheric heating-produced ELF/VLF waves over 32 hours, *J. Geophys. Res., 92*, 8707, 1987.

[25] Rietveld, M. T., P. Stubbe, and H. Kopka, On the frequency dependence of ELF/VLF waves produced by modulated ionospheric heating, *Radio Sci., 24*, 270, 1989.

[26] Rowland, H. L., Simulations of ELF radiation generated by heating the high-latitude D region, *J. Geophys. Res., 104*, 4319, 1999.

[27] Schulz, G. J., Vibrational excitation of N_2, CO, and H_2 by electron impact, *Phys. Rev., 135A*, 988, 1964.

[28] Shkarofsky, I. P., T. W. Johnston, and M. P. Bachynski, The particle kinetics of plasmas, Addison-Wesley, Reading, Massachusetts, 1966.

[29] Stubbe, P., and H. Kopka, Modulation of the polar electrojet by powerful HF waves, *J. Geophys. Res., 82*, 2319, 1977.

[30] Stubbe, P., H. Kopka, and R. L. Dowden, Generation of ELF and VLF waves by polar electrojet modulation: experimental results, *J. Geophys. Res., 86*, 9073, 1981.

[31] Stubbe, P., H. Kopka, M. T. Rietveld, and R. L. Dowden, ELF and VLF wave generation by modulated HF heating of the current carrying lower ionosphere, *J. Atmos. Terr. Phys., 44*, 1123, 1982.

[32] Tsang, K., K. Papadopoulos, A. Drobot, P. Vitello, T. Wallace, and R. Shanny, RF ionization of the lower ionosphere, *Radio Sci., 26*, 1345, 1991.

[33] Yee, K., Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas Propagat., 14*, 302, 1966.

[34] Zhou, H. B., K. Papadopoulos, A. S. Sharma, and C. L. Chang, Electromagneto-hydrodynamic response of a plasma to an external current pulse, *Phys. Plasmas, 3*, 1484, 1996.