Numerical Simulation of Stimulated Electromagnetic Emissions in the Ionosphere

by

K.T. Cheng

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering.

Approved by :

W. A. Scales

Dr. W.A.Scales, Chairman

Gary S. Brown

Dr. G.S. Brown

Dr. I.M. Besieris

July 1993
Blacksburg, VA

Abstract

One-dimensional electrostatic plasma simulation using the particle-in-cell technique is used to study the spectral features of stimulated electromagnetic emission (SEE). SEE is a potential diagnostic tool to study the ionosphere and its spectrum provides a different view of the heated region from the incoherent scatter radar. At this time, a unified and complete theory which explains the SEE phenomena in detail does not exist. The SEE simulations we discuss are proposed to provide interpretation of many of the past puzzles of the experimental data, as well as to facilitate the design of future SEE experiments and the theoretical development of SEE.

In the numerical simulation, only the upper hybrid layer where the geomagnetic field is essentially perpendicular to the density gradient is modelled. Three of the SEE features, namely the downshifted maximum (DM), upshifted maximum (UM) and broad upshifted maximum (BUM), are suggested to be generated at the upper hybrid layer. We observed these three features which have many similarities in the simulation. It is evident that the DM and UM are generated by the same parametric instability involving lower hybrid waves while the BUM is produced by other different mechanisms. Boundary effects are found important on the generation of all three features in the simulation. Moreover, detailed investigation of the simulation results raises a number of questions concerning detailed generation mechanisms of SEE which have not been considered and answered in the past.

Besides the DM, UM and BUM features, the quenching of DM is also observed in the simulation when the pump frequency is very close to electron cyclotron harmonics. It is concluded that both the cyclotron damping and mode conversion of the upper hybrid wave into electron Bernstein modes are possible causes. Finally, some suggestions for the future SEE simulation are included.

Acknowledgement

I wish to express my gratitude to my advisor, Professor W.A. Scales, for his advice, aid and guidance during my master study. I am especially grateful to him for his untired effort and enthusiasm in teaching me the basic plasma theory and in the discussion of many problems I encountered during my research. Without his support and encouragement, hardly could this thesis be accomplished in one year.

I also wish to thank Professors G.S. Brown and I.M. Besieris for serving on my master advisory committee. In particular, I am indebted to Professor Brown for his enthusiastic introduction a year ago so that I could have an opportunity to work with my advisor.

Finally, I would like to acknowledge my thanks to my parents for their support and patience during my master study in the United States.

This work has been supported by Office of Naval Research grant N00014-92-J-1498.

Index of figures

Figure 1: Production of ionization in the atmosphere. p.5

Figure 2: Sample ionograms taken on a summer daytime with O-wave only. p.8

Figure 3: (a) A typical ISR spectrum of F region, with a radar frequency of 430MHz.
(b) ISR spectra at different height in the ionosphere. p.9

Figure 4: Detailed electron density profile from 100km to 300km. p.12

Figure 5: Electron density and temperature profiles of the ionosphere. p.13

Figure 6: Height variations of the main atmospheric constituents. p.14

Figure 7: Dependence of the effective electron collision frequency upon height. p.15

Figure 8: Primary SEE features. p.18

Figure 9: (a) SEE spectrum to demonstrate higher order DMs.
(b) SEE spectrum to demonstrate $\frac{1}{2}DM$ and $\frac{1}{2}UM$. p.20

Figure 10: Five consecutive SEE spectra recorded at Arecibo. p.21

Figure 11: (a) Dispersion of an unmagnetized plasma.
(b) Dispersion of a magnetized plasma with $Y < 1$. p.42

Figure 12: Ray trajectories for the illustration of mode conversion at critical angles. p.44

Figure 13: Hierarchy of heater thresholds in the ionospheric modification experiments. p.51

Figure 14: Sketch of the dispersion curves for electron Bernstein modes. p.58

Figure 15: Illustration of cyclotron damping at $\omega = 2\Omega$. p.60

Figure 16: Discretization of the plasma length and naming of grids and particles in ES1. p.68

Figure 17: Basic algorithm of ES1. p.68

Figure 18: A physical picture of 1D 3V electrostatic plasma simulation. p.69

Figure 19: Leap-frog integration method used in ES1. p.71

Figure 20: Reflective boundary condition used in ES1G. p.73

Figure 21: Linear density gradient used in ES1G program. p.76

Figure 22: ES1G simulation without pump field (noise case). p.87-88

Figure 23: Simulation results at $\omega_o = 0.88$, with $ecenter = ngpsd = 154$. p.91-92

Figure 24: Simulation results at $\omega_o = 0.86$, with $ecenter = ngpsd = 136$. p.95

Figure 25: Simulation results at $\omega_o = 0.84$, with $ecenter = ngpsd = 119$. p.96

Figure 26: Simulation results at $\omega_o = 0.89$, with $ecenter = ngpsd = 163$. p.98-99

Figure 27: Simulation results at $\omega_o = 0.88$, with $lll = 106.8$, $\omega_{peo} = 0.9$, $ecenter = ngpsd = 393$ and $dfn = 1.314$. p.100-101

Figure 28: Simulation results at $\omega_o = 1.18$, with $ecenter = ngpsd = 463$. p.102-103

Figure 29: Simulation results at $\omega_o = 1.16$, with $ecenter = ngpsd = 440$. p.104

Figure 30: Simulation results at $\omega_o = 1.14$, with $ecenter = ngpsd = 417$. p.105

Figure 31: Simulation results at $\omega_o = 1.19$, with $ecenter = ngpsd = 475$. p.106

Figure 32: Simulation results at $\omega_o = 1.195$, with $ecenter = ngpsd = 481$. p.107

Figure 33: Simulation results at $\omega_o = 1.18$ for two different excitation positions. Both spectra are taken from $t = 1000$ to $t = 2600$. p.108

Figure 34: Simulation results at $\omega_o = 0.58$, with $ecenter = ngpsd = 107$. p.109

Figure 35: Simulation results at $\omega_o = 1.48$, with $ecenter = ngpsd = 373$. p.110-111

Figure 36: Simulation result at $\omega_o = 0.88$, with $ecenter = ngpsd = 157$, $K = 100$. p.113

Figure 37: Effect of mass ratio on power spectrum ($\omega_o = 0.88$, $ecenter = ngpsd = 159$, $K = 225$ and Region II parameters). p.113

Figure 38: Effect of observation point on power spectrum ($\omega_o = 0.88$, $ecenter = 154$ and Region II parameters). p.114-115

Figure 39: Effect of observation point on power spectrum ($\omega_o = 0.88$, $ecenter = 393$ and parameters similar to the case in Figure 27). p.116

Figure 40: Effect of electron temperature on power spectrum ($\omega_o = 0.88$, $ecenter = 154$, $v_{te} = 0.5$ and Region II parameters). p.117-118

Figure 41: Effect of electron temperature on power spectrum ($\omega_o = 0.88$, $ecenter = 154$, $v_{te} = 2$ and Region II parameters). p.118

Figure 42: Effect of ion temperature on power spectrum ($\omega_o = 0.88$, $ecenter = 154$, $v_{ti} = 0.0357$ and Region II parameters). p.119

Figure 43: Simulation results at $\omega_o = 0.9$, with $ecenter = ngpsd = 172$. p.121-122

Figure 44: Simulation results at $\omega_o = 1.2$, with $ecenter = ngpsd = 487$. p.123-124

Figure 45: Simulation results at $\omega_o = 0.91$, with $ecenter = ngpsd = 181$. p.125-126

Figure 46: Simulation results at $\omega_o = 0.92$, with $ecenter = ngpsd = 190$. p.128

Figure 47: Simulation results at $\omega_o = 0.94$, with $ecenter = ngpsd = 209$. p.129

Figure 48: Simulation results at $\omega_o = 0.96$, with $ecenter = ngpsd = 228$. p.130

Figure 49: Simulation results at $\omega_o = 0.905$, with $ecenter = ngpsd = 176$. p.131

Figure 50: Simulation results at $\omega_o = 1.22$, with $ecenter = ngpsd = 511$. p.133-134

Figure 51: Simulation results at $\omega_o = 1.24$, with $ecenter = ngpsd = 536$. p.135

Figure 52: Simulation results at $\omega_o = 1.26$, with $ecenter = ngpsd = 561$. p.136-137

Figure 53: Simulation results at $\omega_o = 1.205$, with $ecenter = ngpsd = 493$. p.138

Figure 54: Simulation results at $\omega_o = 1.21$, with $ecenter = ngpsd = 499$. p.139

Figure 55: Simulation results at $\omega_o = 1.205$, with $ecenter = ngpsd = 177$ and other parameters similar to those shown in Figure 33(a). p.140

Figure 56: Simulation results at $\omega_o = 1.22$, with $ecenter = ngpsd = 773$ and other parameters similar to those shown in Figure 33(b). p.142

Figure 57: Simulation results at $\omega_o = 0.91$, with $ecenter = ngpsd = 186$, $K = 225$. p.143

Figure 58: Simulation results at $\omega_o = 0.91$, with $ecenter = ngpsd = 181$, $v_{te} = 0.5$. p.144-145

Figure 59: Simulation results at $\omega_o = 0.91$, with $ecenter = ngpsd = 181$, $v_{te} = 2$. p.146

Figure 60: Simulation results at $\omega_o = 0.91$, with $ecenter = ngpsd = 181$, $v_{ti} = 0.0357$. p.147

Figure 61: Simulation results at $\omega_o = 1.125$, with $ecenter = ngpsd = 107$. p.148

Index of tables

Table 1: Summary of the parameters in the three regions in Figure 4. p.11

Table 2: Summary of the heater characteristics in Arecibo and Tromsϕ. p.16

Table 3: Parameter sets used in ES1G simulation for the three defined regions. p.85

Table of contents

1. Introduction and objectives p.1

2. Ionosphere and stimulated electromagnetic emissions (SEE) p.3

2.1 Formation of ionosphere p.4

2.2 Radio sounding of the ionosphere p.7

2.3 Physical parameters of the ionosphere p.11

2.4 Ionospheric modification experiments and SEE p.13

3. Basic plasma theory p.22

3.1 Electrostatic electron plasma oscillation (Langmuir wave) p.22

3.2 Gyromotion of a charged particle in uniform magnetic field p.23

3.3 Debye shielding p.24

3.4 Dielectric tensor and wave propagation in a cold magnetized plasma p.26

3.5 Kinetic and fluid descriptions of a plasma p.34

3.5.1 Kinetic description (Vlasov Equation) p.35

3.5.2 Plasma as fluids p.36

3.6 Mode conversion p.40

3.6.1 Linear mode conversion p.41

3.6.2 Direct conversion p.45

3.7 Pondermotive force and parametric instability p.46

3.8 A survey of the proposed SEE theories p.51

4. 1-D numerical simulation p.65

4.1 Types of numerical simulation of plasmas p.65

4.2 1-D particle-in-cell electrostatic plasma simulation p.67

4.2.1 A physical picture p.69

4.2.2 Equation of motion (Lorentz force equation) p.69

4.2.3 Field equation (Poisson's Equation) p.72

4.2.4 Particle position and velocity loading with density gradient p.75

4.2.5 Precautions about numerical accuracy p.78

4.2.6 Diagnostics p.79

5. Numerical simulation of SEE in the ionosphere p.80

5.1 Calulation of power spectrum p.80

5.2 Simulation setup p.83

5.3 Simulation results p.89

5.4 Discussions p.144

6. Summary and conclusion p.153

Appendix

(A) INIT subroutine listing p.155

(B) PSD program listing p.159

(C) Input/output file formats and examples p.171

(D) List of simulated cases p.187

(E) UH program listing p.189

(F) HDFCONV program listing p.191

References p.193

Vita p.197

1 Introduction and objectives

The interaction of electromagnetic waves with the ionosphere has been an active research area since the landmark trans-Atlantic experiment by Marconi in 1901. Evidence of modification of the ionosphere by a powerful radio wave was first observed in 1933 when a powerful transmitting station in Luxemborg was found to modulate signals transmitted from Switzerland to Holland. Bailey and Martyn proposed that the signal from the Luxemberg station had increased the ionosphere electron temperature and thus had modulated the radio wave absorption. Since that time, there has been growing interest in "heating" the ionosphere with high power radio waves.

In the past two decades, a number of heating facilities have been built in Europe, Russia and the United States to study the remote interaction of high frequency electromagnetic waves with the ionospheric plasma. Modern technology has resulted in the development of powerful transmitters that can produce strong modifications to the ionosphere. Much theoretical, numerical and experimental work has been performed during this time period. A large body of work was produced in the early 70s at newly built U.S. facilities in Platteville, Colorado and Arecibo, Puerto Rico. Some of this early pioneering research is described in the November 1974 review issue of Radio Science. Much of the theoretical work performed in the Soviet Union during this time period is reported in the two Russian monographs by Ginzburg (1970) and Gurivech (1978), which have English versions. More recent results of work performed during the 80s at heating facilities in Tromsϕ, Norway, Arecibo, Purto Rico and the HIPAS facility in Fairbanks, Alaska, is summarized in the November 1990 special review issue of Radio Science. Current topics of interest in ionospheric modification research include such diverse aspects as modification of the polar electrojet and ELF/VLF generation, hot electrons and artificial airglow emissions, large scale density and temperature modifications and generation of electrostatic waves, parametric instabilities and production of small-scale density irregularities.

Another topic of current interest in ionospheric modification is stimulated electromagnetic emissions (SEE). During heating experiments near Tromsϕ, Norway, it was discovered that when a powerful O-mode electromagnetic pump wave, which has a frequency near the harmonics of electron cyclotron frequency, is injected into the ionosphere from a ground station, secondary electromagnetic waves are generated and can be detected on the ground ([22] Thide 1982 [1]). These electromagnetic waves have frequencies that are in a relatively small bandwidth around the pump wave frequency. Under varying pump wave and ionospheric conditions, these waves may be at frequencies above or below the pump frequency. Since those first experiments, SEE has also been observed at Arecibo

[1] There are two types of references in this thesis. One is text or dissertation, quoted with reference number and the name of the first author. The other is published papers, cited in addition with publication year.

([23] Thide 1990), Alaska ([2] Armstrong 1990) and Russia ([3] Boiko 1985).

The importance of the SEE spectrum has been emphasized as a useful diagnosis for studying parametric instabilities and other nonlinear physical processes that may occur around the heated region in [20] Stubbe 1984. It has been shown by Leyser and Thide ([9] Leyser 1988) that the spectrum from SEE may be used to measure the electric field strength in the heated region. The SEE spectrum was used by Leyser ([10] Leyser 1989) to measure the magnitude of the background magnetic field. These results show that SEE can potentially be an important diagnostic tool which can complement current diagnostics during ionospheric modification experiments such as incoherent scatter radar and ground-based optical measurements. The SEE spectrum provides a different view of the heated region from the incoherent scatter radar (ISR) spectrum since the SEE spectrum is produced by all wave vectors. The ISR spectrum is produced by one wave vector because the radar selects only one wave vector.

It is the goal of this research work to study some of the physical processes which produce SEE in more detail than the past work, mainly by using numerical simulations. As far as the author knows, until now, we do not have a unified and complete theory on all the observed phenomena of SEE. The predictions from this study will provide interpretation of many of the past puzzles of the experimental data, as well as insight into the nonlinear phenomena that can occur during ionospheric heating experiments, the design of future SEE experiments and the use of SEE as a diagnostic tool. Numerical simulations have some distinct advantages over the experiments in some aspects; e.g., one can change any physical parameters in the plasma in a simulation. Ultimately, by the correlation of the experimental work and numerical simulations, we can know more about the SEE generation.

The outline of this thesis is as follows. In Chapter 2, we will briefly discuss the formation, radio sounding and some important parameters of the ionosphere. Also we will introduce the ionospheric modification experiments and SEE, and then summarize most of the SEE experimental data obtained during the last decade. In Chapter 3, we will review the most important and fundamental theories in plasma physics, which are related to our later SEE discussion. Also, at the end of Chapter 3, most of the existing SEE theories will be briefly summarized.

Techniques used in one-dimensional electrostatic plasma simulation is the central topic of Chapter 4. It will provide basic explanation about how we deal with 1-D numerical plasma simulation using the particle-in-cell method. Then we will proceed in Chapter 5 to demonstrate how we use the numerical simulation discussed in the previous chapter to study the SEE. Having the simulation results, we can discuss their correlations with the experimental data. Here, some important points about the SEE generation that may have been neglected in the past work, are addressed. Finally, we will summarize our conclusions in Chapter 6.

2 Ionosphere and stimulated electromagnetic emissions (SEE)

The gross structure of the Earth's atmosphere can be described in several ways, and each leads to a classification of the height regions that is appropriate to the physcial process under consideration. Four commonly used classifications are :

- Classification via temperature (e.g. troposphere, mesosphere, etc.)
- Classification based on chemistry (e.g. ozonsphere)
- Classification based on ionization (e.g. ionosphere, protonosphere, etc.)
- Classification via dynamics (e.g. turbosphere, diffusosphere, etc.)

In this thesis, we will focus on the ionized part of the atmosphere, the ionosphere, which contains significant numbers of free electrons and positive ions. There are also negative ions at lower altitudes. The medium as a whole is electrically neutral. Although the charged particles may be only a minority amongst the neutral ones, they exert a great influence on the medium's electrical properties, and herein lies their importance.

After Marconi transmitted a radio signal from Cornwall in England to Newfoundland in Canada in 1901, Kennelly and Heaviside independently suggested that, because of the Earth's curvature, the waves must have been reflected from an ionized layer. The name ionsphere was coined by R. Watson-Watt in 1926.

Since that time, the ionosphere has been extensively studied and most of its principal features, though not all, are now fairly well understood in terms of the physical and chemical processes of the upper atmosphere. The main regions in the ionosphere are classified as D, E, F1 and F2, with the following daytime characteristics :

Name	Altitudes	Electron density
D region	60-90 Km	$10^8 - 10^{10} m^{-3}$
E region	105-160 Km	several $10^{11} m^{-3}$
F1 region	160-180 Km	several $10^{11} - 10^{12} m^{-3}$
F2 region	maximum variable around 300 Km	up to several $10^{12} m^{-3}$

The D and F1 regions vanish at night, and the E region becomes much weaker. The F2 region, however, tends to persist though at reduced intensity.

In this chapter, we will explore the formation of the ionosphere and the ways we can measure the electron density profile from a ground based station in the following two sections. Then some important numerical data of the electron density, molecule constitution, temperature and collision frequency in the ionosphere are displayed. In the last section, a brief review of ionospheric modification experiments and stimulated electromagnetic emissions (SEE) is given. Also, some of the experimental data of SEE are depicted to illustrate the main features of SEE.

2.1 Formation of ionosphere

The ionosphere is formed by the ionization of atmospheric gases such as N_2, O_2 and O. At middle and low latitude, the energy required comes from solar radiation in the extreme ultra-violet (EUV) and X-ray parts of the spectrum. Once formed, the ions and electrons tend to recombine and to react with other gaseous species to produce other ions. Thus, there is a dynamic equilibrium in which the net concentration of free electrons (electron density) depends on the relative speed of the production and loss processes. In general terms, the rate of change of electron density is expressed by a continuity equation :

$$\frac{\partial N}{\partial t} = q - L - \nabla(N\mathbf{v}) \tag{1}$$

where q is the production rate, L the loss rate by recombination, and the divergence expresses the loss of electrons by movement, **v** being their mean drift velocity.

Let us consider the production first. The ionizing energy is incident from above the atmosphere. It will encounter an increasing concentration of ionizable atmospheric particles as it penetrates and so will produce ionization at an increasing rate. The altitude dependence of the ionizable materials is due to gravity. The ionizing energy will then suffer absorption in the process and this will ultimately offset the increase of atmospheric concentration. A peak rate of production of ionization will be attained at a certain height, and at lower heights, the rate will decline until the flux of ionizing energy becomes negligible. The whole production process is picturally depicted in Figure 1.

The rate of production of ion-electron pairs can be expressed as the product of four terms: $q = \eta\sigma n I$. Here, I is the intensity of ionizing radiation at some level of the atmosphere and n is the concentration of ionizable atoms or molecules. For an atom or molecule to be ionized, it must first absorb radiation, and the amount absorbed is expressed by the absorption cross-section, σ. If the incident radiation is I (J/m^2s), then the total energy absorbed per unit volume of the atmosphere per unit time is $\sigma n I$. However, not all this energy will go into the ionization process, and the ionization efficiency, η, takes that into account. From this simple beginning, S. Chapman, in 1931, developed a formula which predicts the form of a simple ionospheric layer and how it varies during the day. The formula expressing the rate of production q is the Chapman production function. The

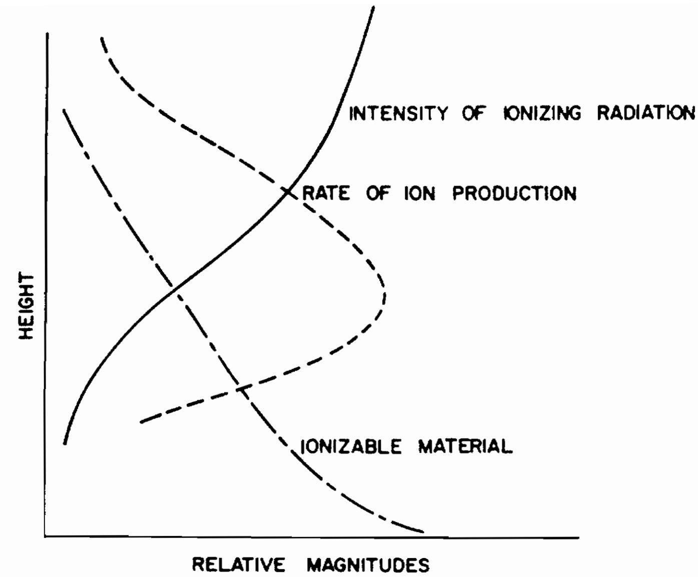


Figure 1: Production of ionization in the atmosphere. (From [6] Hines)

derivation makes the following assumptions :

1. The atmosphere is composed of a single species, exponentially distributed with constant scale height H=kT/mg ([5] Hargreaves).
2. The atmosphere is plane stratified.
3. Solar radiation is absorbed in proportion to the concentration of gas particles.
4. The ionizing radiation is monochromatic so that the absorption coefficient is constant.

The derivation of the Chapman production function can be found in [5] Hargreaves or [4] Davies. The Chapman production function is usually written in a normalized form as,

$$q = q_{mo} e^{1-z-\sec\chi \cdot e^{-z}} \tag{2}$$

Here, z is the reduced height for the neutral gas, $z = (h - h_{mo})/H$. χ is the solar zenith angle and h_{mo} is the height of maximum production rate when the Sun is overhead ($\chi = 0$). By differentiating (2), the maximum production rate is $q_m = \eta I_\infty/(eH \sec\chi)$ at the reduced height $z_m = \ln(\sec\chi)$.

The second term in the continuity equation is the rate of recombination. However, electrons are not normally lost by simple recombination with a positive ion. The recombination rate is about 10^5 times greater with molecular ions than with atomic ions. Hence the atomic (oxygen) ions must be converted to molecular ions before the recombination takes place. Moreover, the original molecular ions are converted into another species of molecular ion. In the lower D region at night, electrons can attach themselves to neutral molecules to form negative ions, which can recombine with positive ions. The following reactions are important in the daytime D region and the E and F regions :

$$\begin{array}{llllllll}
O_2^+ & + & e^- & \longrightarrow & O & + & O & + \; 6.96eV \\
N_2^+ & + & e^- & \longrightarrow & N & + & N & + \; 5.82eV \\
NO^+ & + & e^- & \longrightarrow & N & + & O & + \; 2.76eV \\
O^+ & + & O_2 & \longrightarrow & O_2^+ & + & O & + \; 1.53eV \\
O^+ & + & N_2 & \longrightarrow & NO^+ & + & N & + \; 1.09eV
\end{array}$$

The first three reactions are called dissociative recombination reactions; the last two are called charge-transfer or atom-ion exchange reactions or charge- exchange reactions. In the E and F1 layers, the reactions are fast. In the E region, where almost all the ions are molecules, the rate of recombination of electrons equals the rate of recombination of the molecular ions. For such a case, the recombination rate is proportional to the square of electron density, i.e. $L = \alpha N_e^2$. If we neglect the drift of electrons, at equilibrium, we have $q = \alpha N_e^2$. Taking the production rate q from the Chapman production function, we obtain,

$$N_e = N_{eo} e^{(1-z-\sec\chi \cdot e^{-z})/2} \tag{3}$$

and the maximum electron density is $N_m = N_{eo}\sqrt{\cos\chi}$. The peak corresponds the F1 peak. A layer with these properties is sometimes called an α-Chapman layer.

At higher levels of the ionosphere where photoionization of O is dominant, the electron loss is a two-step process (see the above chemical equations). Then the electron loss rate depends linearly on N_e because the formulation of molecular ions is slow and controls the overall rate. Thus, $L = \beta N_e$. At equilibrium,

$$N_e = N_{eo} e^{(1-z-\sec\chi \cdot e^{-z})} \tag{4}$$

However, β is proportional to neutral molecular density because the recombination process involves neutral particles. Therefore, β is expected to vary with height. This profile has no maximum except at very great heights where there is no sufficient number of ionizable molecules. This layer is called a β-Chapman layer.

It is more complicated than the simple theory of the two types of Chapman layers to explain the F2 peak because neither the α-type nor the β-type recombination rate can dominate the overall process. One has to take into account of the concentration of N_2 and O and the diffusion of electrons. Interested readers may refer to [5] Hargreaves.

2.2 Radio sounding of the ionosphere

There are a wide variety of ways to measure the electron density profile in the ionosphere. Radio sounding techniques are known to be widely used. On the other hand, rockets and satellites can make in-situ measurements and they can measure small-scale density irregularities in both vertical and horizontal extents. In this section, we will briefly discuss the ionospheric sounding methods because this will naturally lead to the ionospheric modification experiments and then the SEE.

To detect the electron density by radio wave, we rely on two fundamental principles: reflection from the ionosphere and scattering from charged particles.

Ionosonde

One of the oldest and still one of the most important techniques of ionospheric study, is the ionosonde which uses the first principles ([4] Davies). It transmits a radio pulse vertically and measures the time which elapses before the echo is received. Actually, this is a sweep-frequency pulsed radar. The frequency can range from below 0.1MHz to 30MHz with a sweep duration from a few seconds to a few minutes. As will be seen in Chapter 3, the reflection height in the ionosphere depends on the polarizaton of the launched radio wave because left-handed and right-handed circularly polarized waves have different dispersion relations. For the ordinary wave [1] or O-wave, the reflection point is at the height where the plasma frequency equals the pump frequency. The dependence of plasma frequency upon the electron density is discussed in section 3.1. Here we cite the formula as $f_p \approx 9\sqrt{n}$. Then from the ionogram, which is a chart recording the virtual height versus frequency, we can estimate the electron profile. The measured height is virtual because the radio wave passes through a medium with varying refractive index, especially in the vicinity of the reflection. It effectively changes the group velocity of the wave and hence the measured time. Some correction is needed to produce a real-height profile.

Figure 2 shows a sample of conventional ionograms. We can clearly see the E, F1 and F2 regions, and these special signatures on the ionograms in fact gave their names. The three regions are divided because there are three local maxima, two of them denoted by foF1 and foF2 in Figure 2, between these three regions. Beyond the critical frequency [2] foF2, the ionosonde fails to measure the topside electron density because the electromagnetic wave penetrates the ionosphere. Note that the rapidly increasing virtual heights near the frequencies foF1 and foF2 are due to the retardation of the group

[1] In the northern hemisphere, it is right-handed polarized with respect to its wave vector.

[2] It occurs at the height where the electron density is maximum. Then the corresponding plasma frequency is called the critical frequency of the ionosphere.

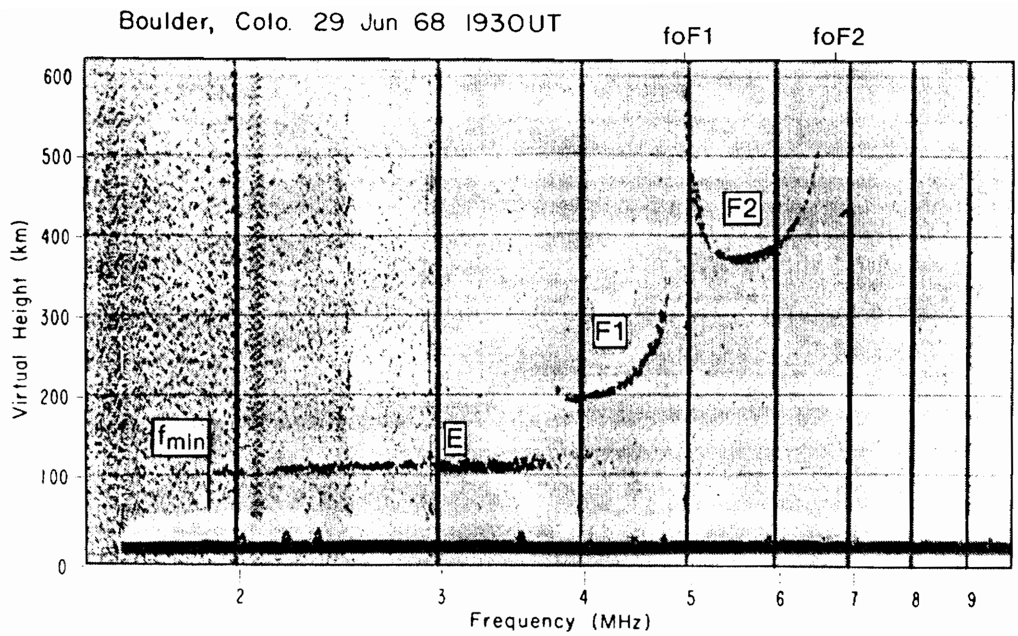


Figure 2: Sample ionograms taken on a summer daytime with O-wave only (From [4] Davies)

velocities of the sounding wave. The topside measurement has to rely on a satellite carrying a topside sounder to give us information (e.g. Alouette I, [6] Hines).

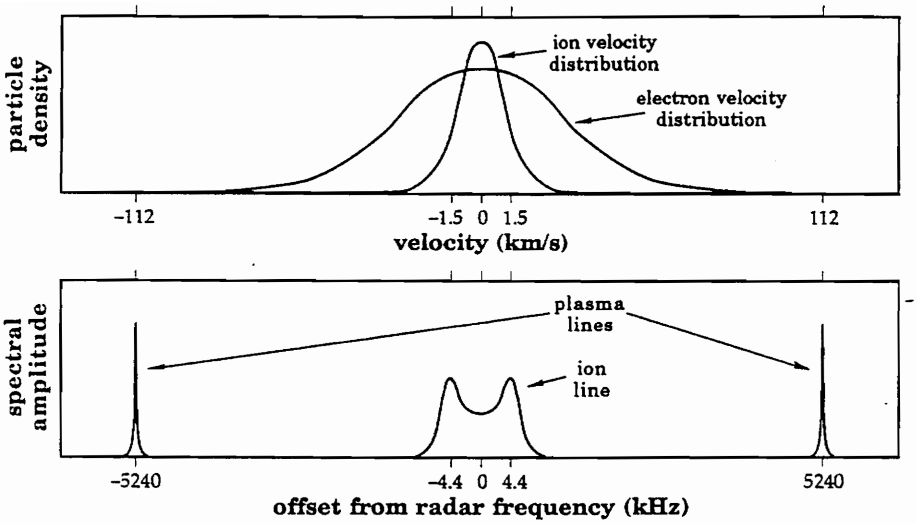
There were a number of improvements of the conventional ionosondes, such as the chirp technique ([4] Davies). However, two major disadvantages remove the ionosondes from an ideal instrument of ionospheric sounding. They are the following:

- It has no way to give direct information on "valleys" between layers.
- Its resolution is not good enough to measure the fine details of the electron density profile.

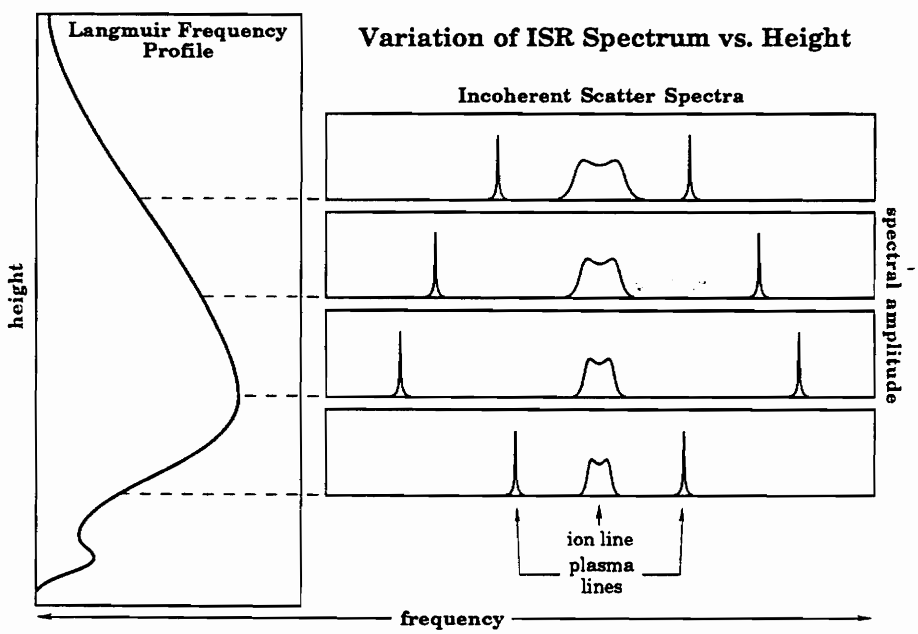
Incoherent scatter radar

Although there are some alternatives to the ionosondes (e.g. Doppler sounder) using the reflection principle, the incoherent scatter radar (ISR), however, is proved to be successful in radio sounding of the ionosphere. It makes possible not only the measurement of the topside ionosphere from the ground but also the measurement of a variety of other properties of the upper atmosphere (neutral density, temperature and composition).

In 1906, Thomson showed that electrons are capable of scattering electromagnetic waves (X-rays). The scattering cross section ($\sigma_e = 4\pi r_e^2 = 10^{-28}\ m^2$) of an electron is independent of the wavelength λ and is called Thomson cross section. Gordon in 1958 proposed that a radar could detect weak scatter from the ionosphere and it was achieved in practice by Bowles in the same year. Gordon



(a)



(b)

Figure 3: (a) A typical ISR spectrum of F region, with a radar frequency of 430MHz. (b) ISR spectra at different height in the ionosphere. (Both from [8] Isham)

predicted that the random thermal motions of the electrons would produce Gaussian broadening of the spectrum with a center to half power width of $0.71\Delta f_e$, where Δf_e is the Doppler shift produced by an electron approaching the radar at the mean thermal speed $v_t = (2k_B T_e/m_e)^{1/2}$. Thus, one would expect,

$$\Delta f_e = \frac{2\sqrt{2k_B T_e/m_e}}{\lambda} \approx 11\frac{\sqrt{T_e}}{\lambda} \quad \text{(in KHz)} \tag{5}$$

The factor two is due to electrons moving towards and away from the radar. For a wavelength of 0.75m and $T_e = 1600K$, $\Delta f_e \approx 600KHz$. The echo spectrum is broad. However, the actual spectrum observed is some 200 times narrower than this because the motion of the electrons is controlled by the ions. Heavy ions affect radio waves when the probing wavelength is very much longer than the so-called electron Debye length [3] which is about 1cm or less below 1000km, rising to 6cm at 2000km. Each ion may be considered to influence electron motions within a Debye length. Therefore, the existence of heavy ions reduce the spectral width approximately by a factor of $(m_i/m_e)^{1/2}$, where m_i and m_e are the ion and electron masses. The resulting narrower spectrum is called an ion line which has a Doppler shift of,

$$\Delta f_i = \frac{\sqrt{8k_B T_i/m_i}}{\lambda} = 65\frac{\sqrt{T_i}}{\lambda} \quad \text{(in Hz)} \tag{6}$$

For O^+ ions, if $\lambda = 0.75m$ and $T_i = 1600K$, $\Delta f_i \approx 3500Hz$.

Figure 3(a) shows a typical ISR spectrum. It includes two major components — the ion line and plasma line. The evolution of the ion lines has been discussed. The high-frequency plasma lines are due to thermal plasma oscillation of the electrons. The spectrum is nearly symmetrical about the radar frequency, with the upshifted half due to waves traveling towards the radar and the downshifted half due to waves away from the radar. Figure 3(b) illustrates how to measure electron density and molecule composition using ISR. The electron density can be calculated from the frequency shift of the plasma lines, while the average ion mass can be estimated from the frequency offset of the ion lines. The scattered signals from the bottomside and topside of the ionosphere are selected with a range gate. We can also estimate the thermal velocities or temperatures of both electrons and ions from ISR spectrum.

Since the first use of ISR in late 50s, there were a large number of improvements and subsequent developments of the incoherent scatter radar (e.g. chirped ISR, [8] Isham). However, one major disadvantage of ISR is that one has to work with very weak signal. This limits the ISR ability to make high resolution measurements.

[3] The Debye length will be discussed in section 3.3.

2.3 Physical parameters of the ionosphere

Before we go into ionospheric modification and stimulated electromagnetic emissions, we explore some of the important parameters of the ionosphere. These are valuable references for the interpretation of the SEE data and for the setup of the numerical simulation.

The first parameter is the electron density profile. Although the profile is highly dynamical, we will provide some typical values. Figure 4 shows a daytime electron density profile. The critical frequency is about 7MHz, peaking at 250km. We model the bottomside profile by three linear regions — I, II and III, each of which has end points shown in Figure 4. Region I is approximately the top E-layer; whereas Region II and III are F1 and bottom F2 layers, respectively. We will use the normalized density difference (dfn[4]) in each region to set up our numerical simulations in Chapter 5. Table 1 summarizes the parameters in these three regions.

Table 1: Summary of the parameters in the three regions in Figure 4.

Region	Height range	Density range	Slope	dfn	Reference density/frequency
I	106.3,196.6	0.5,2.5	0.02216	1.333	1.5/3.47
II	169.7,214.7	1.2,5.8	0.1023	1.314	3.5/5.30
III	177.7,248.2	3.8,6.5	0.03829	0.330	5.15/6.43
	km	$\times 10^{11} m^{-3}$	$\times 10^{11} m^{-3}/km$		$\times 10^{11} m^{-3}/MHz$

Figure 5 shows a comprehensive graph of the variations of the electron density and temperature against height. In the density profiles, it clearly indicates where the D, E, F1 and F2 regions are, and at nighttime, as stated in the beginning of this chapter, the D and F1 regions disappear. In fact, the density profile may be totally different from those in Figure 5 for both daytime and nighttime. Compared with Figure 5 with Figure 4, one may ask why the F1 peak cannot be found in Figure 4. The reason is simple. It is because either the outdated ionosonde could not pick out the weak peak or the F1 peak was absent when measurement was taking. These curves are used as illustration of the most important features in the ionosphere, but do not serve as standard reference.

The temperature increases rapidly above 120km and reaches approximately 1000K at 200km. Then the profile becomes flat again. Therefore, we deal with hot plasma in the F region. The thermal velocity of electrons is about several hundred km/s and relativistic effect will not be important. Also, the ion temperature is only a quarter of that of electrons in the E and F regions.

The molecular constitution of the atmosphere is depicted in Figure 6. In the E and F layers, oxygen atoms, and hence O^+ ions dominate over all other constituents. This implies to us which

[4] It will be defined in section 4.2.4.

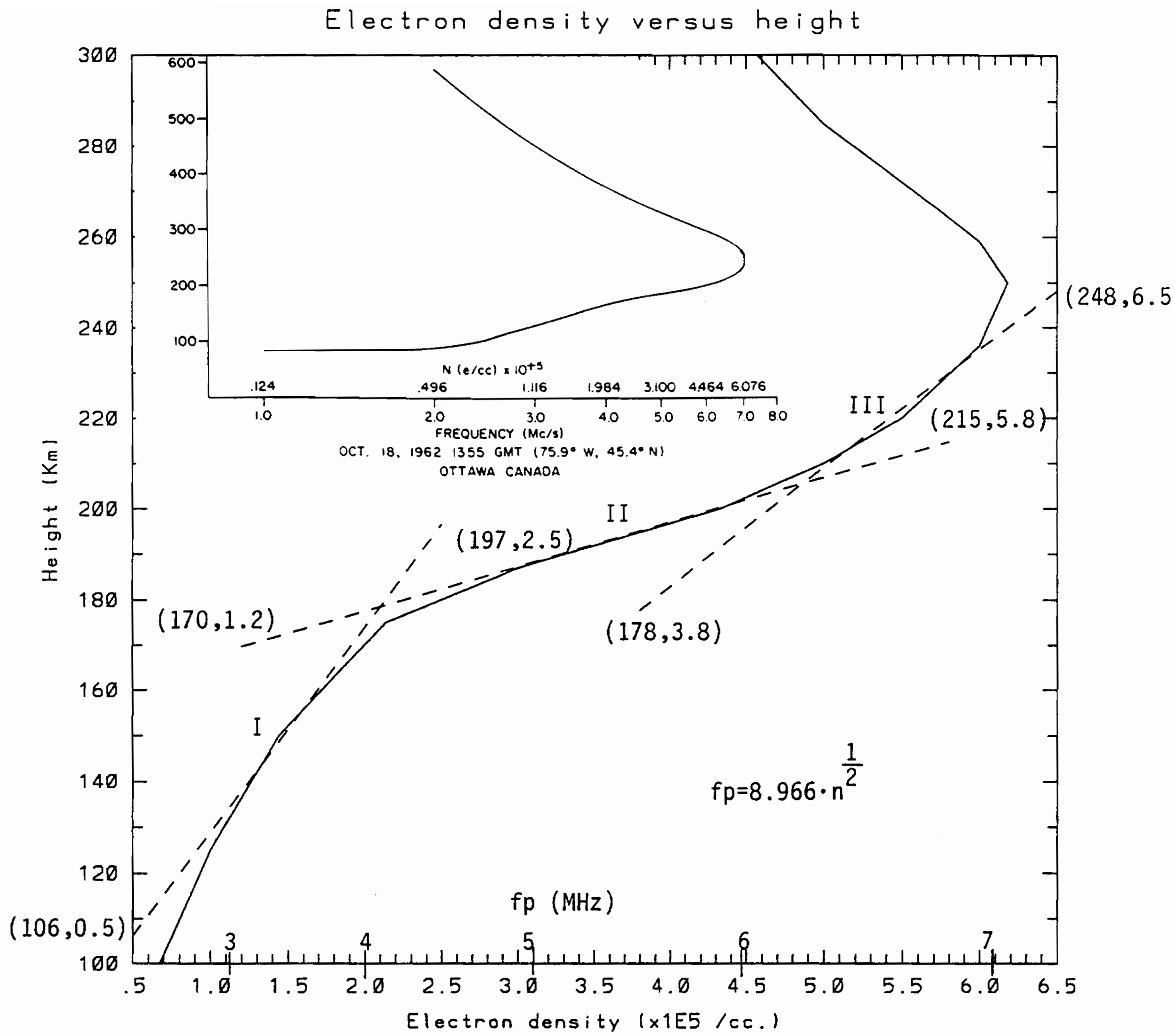


Figure 4: Detailed electron density profile from 100km to 300km. The picture at the left corner is the original profile taken from [6] Hines. The center profile was manipulated so that the density axis is linear. The six ordered pairs are end-point values of the three regions.

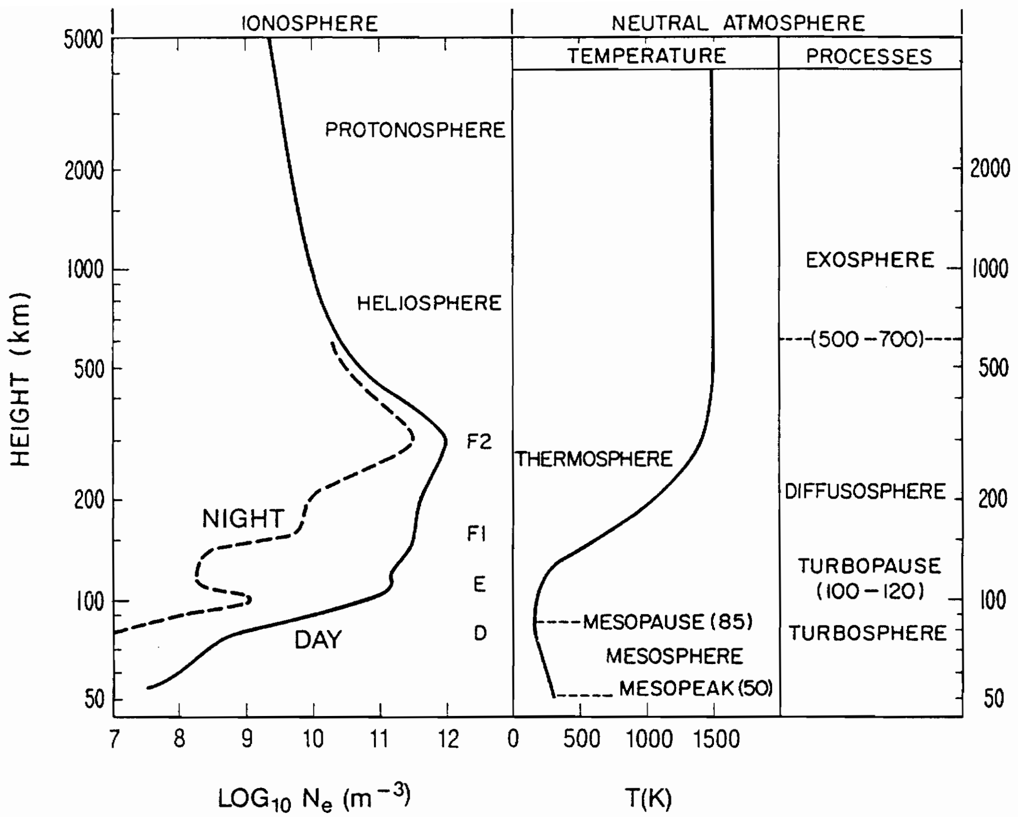


Figure 5: Electron density profiles at daytime and nighttime, and temperature profile versus height. (From [4] Davies)

mass ratio we should use to calculate parameters involving ion-electron interaction in the ionosphere. Also, compared Figure 6 with Figure 4 or 5, one would discover the neutral oxygen atom is about four orders of magnitude higher than the O^+ ions and electron densities. So we may worry about the collision between charged and neutral particles. Figure 7 shows the collision frequency against height. Note that the collision frequency in the F region is far below the plasma frequency in which we are interested.

2.4 Ionospheric modification experiments and SEE

The purpose of an ionosonde is to measure the reflection height and hence the electron density. We would assume the wave energy does not disturb the electron density profile significantly. But if one increases the pump power so that it may be large enough to substantially change the profile, we will call this the ionospheric modification experiment. In fact, besides using high pump wave

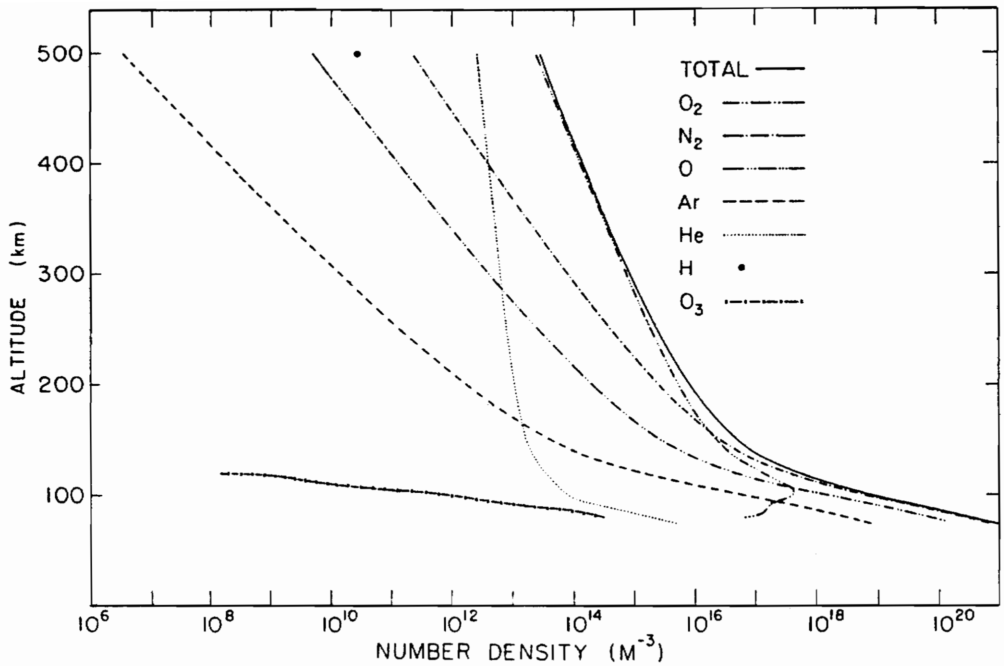


Figure 6: Height variations of the main atmospheric constituents. (From [4] Davies)

power, several other ways (e.g. chemical release, particle beam injections) exist for the ionospheric modification experiments.

Soon after the Marconi trans-Atlantic experiment, people tried to increase the transmitter power. But the ionosphere responds nonlinearly with the incident power. The received signal strength is not directly proportional to the transmitter power.

The ability of ground-based radio transmitters to modify the ionosphere became apparent after the discovery of the ionospheric cross modulation or Luxembourg effect in 1933. In 1938, Bailey advanced the idea of ionospheric modification via gyro-heating. The primary objectives of early ionospheric heating [5] are two-fold :

- To artificially increase the local electron density to facilitate the sky wave communication, especially at nighttime.

- To sustain a glow discharge or airglow and thereby visible brighten the night sky.

Although airglow enhancement or density modification by resonant gyro-interaction has yet to be achieved, experiments clearly demonstrated the absorption of modest HF fluxes can significantly raise the electron temperature throughout the bottomside ionosphere.

[5] Ionospheric modification by high power radio waves is often referred to as ionospheric heating.

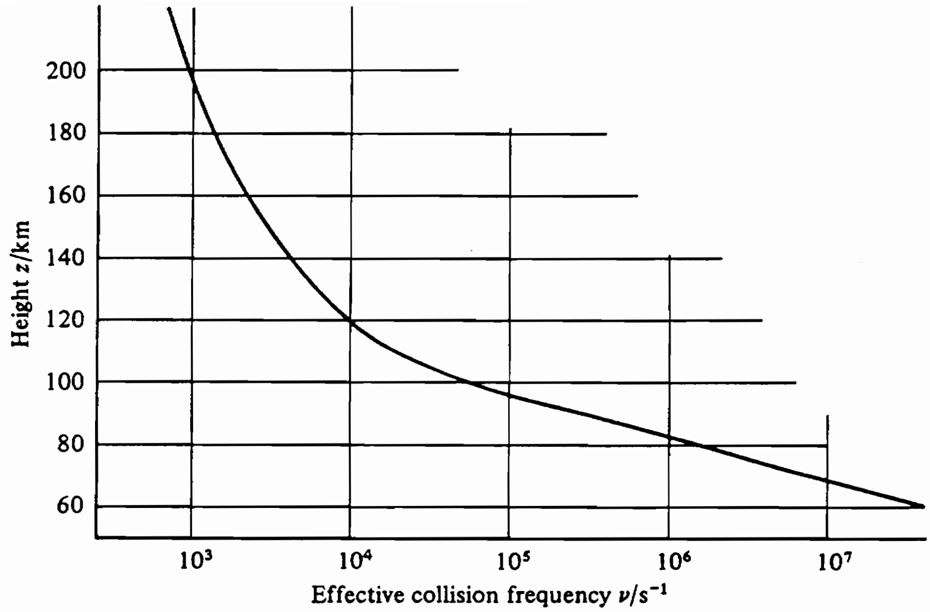


Figure 7: Dependence of the effective electron collision frequency upon height. (From [2] Budden)

Later, people carried out the modification of the E and F regions of the ionosphere through the excitation of the plasma frequency resonance by HF radio waves, referred as ohmic heating. Two commonly used heating frequencies are 3.15 and 5.1MHz. The early theory predicted the observed large-scale changes in the temperature and density, but it left many other phenomena unexplained. These include :

- Anomalous absorption of energy due to the possible excitation of parametric instabilities.
- Airglow and fast electrons.
- Short-scale ($\sim 3m$) geomagnetic field-aligned density irregularities.
- Large-scale ($\sim 200m$) geomagnetic field-aligned striations (also referred to as spread F).
- Changes in ISR spectrum during heating. They include the parametric decay instability line, the oscillating two-stream instability line, and the broad pump.
- Plasma line asymmetry in ISR spectrum.
- Initial overshoot in ISR echo.

More details about the above phenomena can be found in [8] Isham. Some of them are due to short time-scale processes ($\sim$ msec) and some due to long time-scale processes ($\sim$ seconds or minutes).

Table 2: Summary of the heater characteristics in Arecibo and Tromsϕ (From [8] Isham).

Location	Latitude	Longitude	Magnetic dip	f_{ce}	Heater power	ERP	Frequency range
Arecibo	18.4°N	66.8°W	48.5°	0.98	0.4	80	2.5-18
Tromsϕ	69.6°N	19.2°E	77°	1.35	1.5	360	3.85-8
				MHz	MW	MW	MHz

These surprising observations triggered a series of systematic ionospheric heating experiments in the 60s and 70s. The ionosphere was then recognized as a natural laboratory for plasma physics, which virtually has no boundary. Since then, a number of heating facilities were built all over the world. In the United States, heaters were built in Platteville, Colorado and HIPAS, Alaska. The Platteville heater, however, is not operating today. The HIPAS heater will be upgraded. There are several other ionospheric heaters around the world (Norway, Puerto Rico and Russia). The ones in Tromsϕ, Norway, and Arecibo, Puerto Rico, are the two important heater stations to supply us the experimental data of the stimulation electromagnetic emissions. Table 2 summarizes some of the important characteristics of the heaters in Arecibo and Tromsϕ. The values for the magnetic dip, which is the angle between the magnetic field and the ground, and the cyclotron frequency (will discuss it in the next chapter) f_{ce} are at 300km. The transmitters are composed of several linear antenna arrays so that the transmitted wave can be linearly and circularly polarized. The receiving antenna is isolated from the transmitter by high mountains. The 305-meter reflector antenna at Arecibo shown in the antenna book by Kraus (p.607, Figure 12-48, 1988 edition) is used for receiving return signals from both ISR and stimulated electromagnetic emission experiments ([11] Leyser).

In 1981, during F region heating experiments near Tromsϕ, Thide *et al* discovered that when a powerful electromagnetic wave is lauched vertically into the ionosphere, regular sideband structures were observed by monitoring the emissions directly on the ground with a spectrum analyzer connected to a receiving antenna ([22] Thide 1982). This is referred as stimulated electromagnetic emissions (SEE). The strength of the sidebands are profound when the pump frequency steps around the harmonics of the electron cyclotron frequency ($\sim \pm 5\%$). Most of these sidebands have signal strength of at least 10 to 20 dB above background noise level. These waves show up in the received wave power spectrum at frequencies in roughly a 100KHz range around the pump frequency. However, the wave power is typically skewed towards the lower sideband. Typical pump frequencies used during the experiments are between 2 and 7MHz with roughly 100-300MW of effective radiated power (ERP). The spectrum can vary significantly depending on ionospheric and pump wave parameters.

The SEE spectrum exhibits a rich variety of waves that may persist for pump periods of several minutes. Some of the SEE features are prominent, systematic and repeatable, which we call the primary features. Most of them are systematically classified in [20] Stubbe 1984. The displayed experimental spectra and observations are cited from [11] Leyser, [11] Leyser 1990, [13] Leyser 1992, [20] Stubbe 1984, [21] Stubbe 1990, [22] Thide 1982 and [23] Thide 1989. Most of these experiments were performed in Tromsϕ in the last decade. The primary features include the continuum, downshifted peak (DP), downshifted maximum (DM), upshifted maximum (UM), broad upshifted maximum (BUM), broad symmetrical structure (BSS) and quenching of DM. They are shown in Figure 8, which were recorded at Tromsϕ.

In the early SEE experiments, it was observed that SEE features only developed for O-mode excitation ([22] Thide 1982). When the wave polarization was changed from O-mode to X-mode, the sideband structures almost disappeared. A summary of the primary SEE features is provided as follows.

- Continuum — It is a broad asymmetrical structure as shown in Figure 8(a). The continuum is the most commonly observed spectral feature when the pump frequency is below 4MHz. Its frequency coverage is highly variable and can extend to 50KHz below the pump frequency. The development of the continuum (and DP) feature seems to be favored when the ionospheric critical frequency is well above the pump frequency and the ionogram shows unperturbed echoes ([11] Leyser). When the pump frequency is around 4.04MHz (slightly below the third harmonic of the electron cyclotron frequency, $3f_{ce}$), the continuum often shows up with DP. The continuum is present even for very low pump powers (86KW).

- Downshifted peak (DP) — It is a short and narrow peak riding the continuum, as shown in Figure 8(a). The downshifted peak feature was only observed when the pump frequency is very close to $3f_{ce}$ ($\approx 4.08MHz$). Its offset frequency from the pump is about 2KHz. Since the DP feature always shows up with the continuum, the favorate ionospheric condition for the continuum is also applied here.

- Downshifted maximum (DM) — It occurs as a lower sideband with a considerably sharp cutoff on its high frequency side (Figure 8(b)). The cutoff frequency is approximately 7 to 8KHz below the pump frequency, while the bandwidth of the DM feature is usually less than 5KHz. The frequency offset of its peak from the pump is in the range of 8 to 13KHz and increases with pump frequencies ($\Delta f_{DM} \approx 2 \times 10^{-3} f_o$, where f_o is the pump frequency). Its shape is highly variable and tends to skew towards the high frequency side. The DM feature is the most commonly observed spectral feature for pump frequencies above 4MHz. This is in

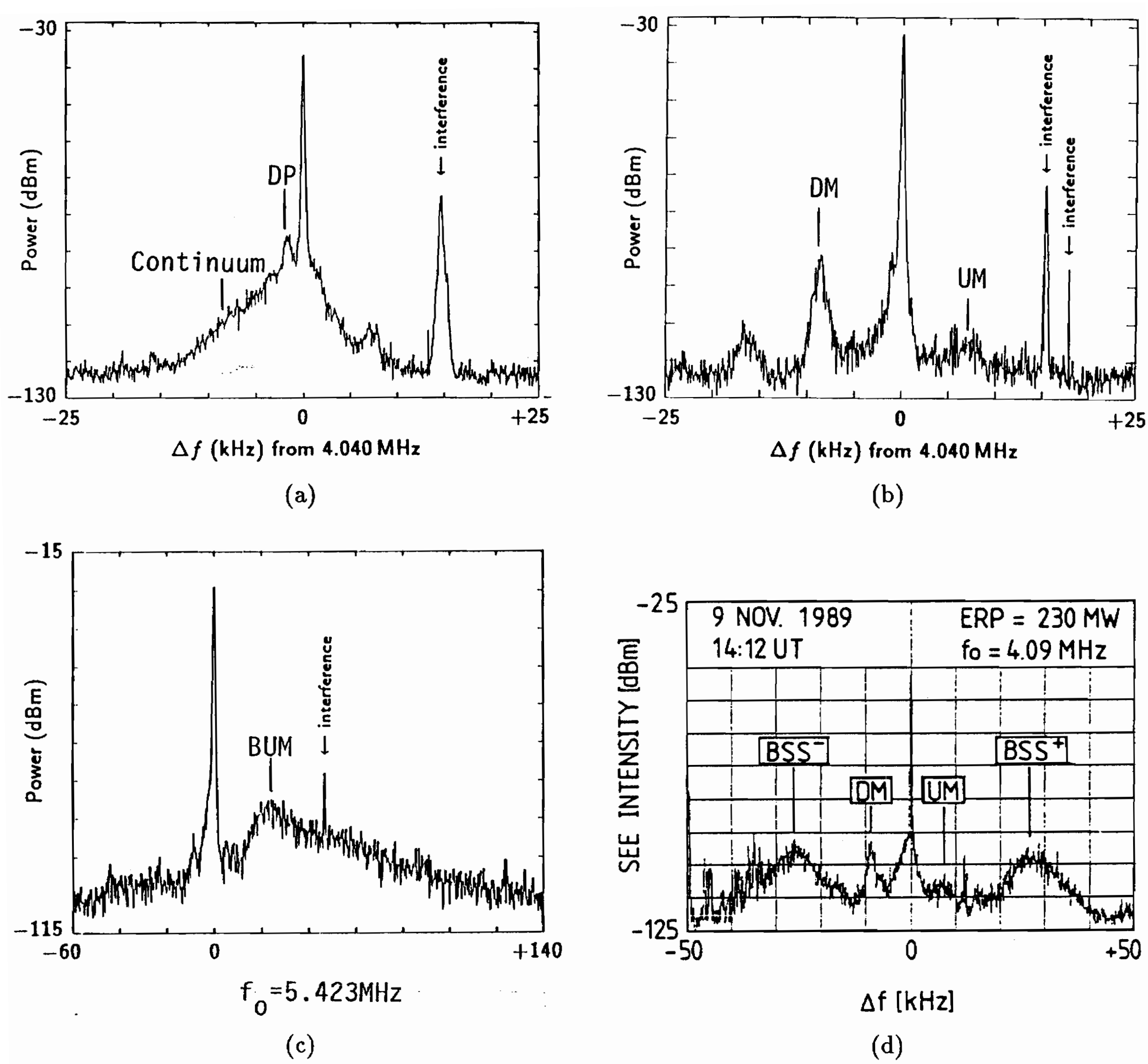


Figure 8: Primary SEE features — (a) Continuum and DP, recorded at 11:20 UT on 27 October 1984; (b) DM and UM, at 9:45 UT on 27 October 1984; (c) BUM, at 15:01 UT on 12 May 1988; (d) BSS, at 14:12 UT on 9 November 1989 ((a)-(c) from [11] Leyser and (d) from [21] Stubbe 1990).

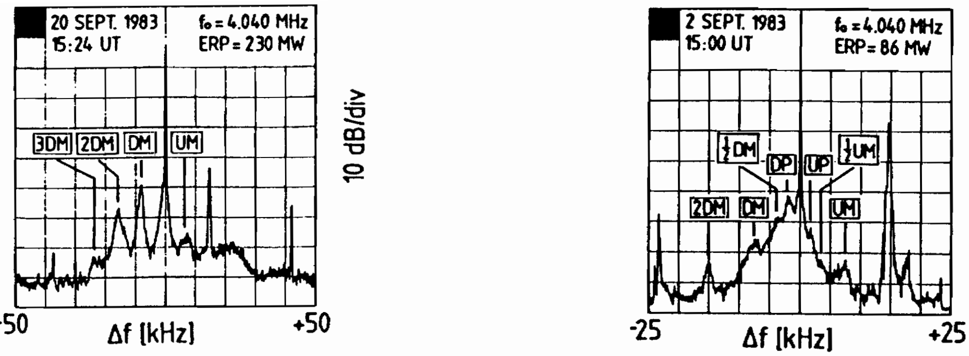
contrast with the continuum feature. The development of DM (and the higher order DMs) is favored when the critical frequency of the ionosphere is near the pump frequency (around 4.04MHz). Experiments showed that the DM developed even when the critical frequency was a few 100KHz below the pump frequency in the morning ([11] Leyser). Also from the experimental observations ([11] Leyser), the SEE are strong whenever ISR echoes are strong, and the polarization of the SEE is primarily in the ordinary mode. However, the DM feature is quenched when the pump frequency is very close to the electron cyclotron harmonics (nf_{ce}, $n = 3, 4, 5, 6, 7$).

- Upshifted maximum (UM) — It appears as a upper sideband, but is a considerably weaker feature than the DM (Figure 8(b)). The frequency of the UM peak is about 5 to 9KHz above the pump frequency, which is approximately 35% less than the DM peak. In most of the cases, it shows up with the DM, but the reverse is not always true.

- Broad upshifted maximum (BUM) — It is a broad spectral feature which appears at frequencies higher than the pump (Figure 8(c)). It develops only when the pump frequency is near but larger than the third, the fourth, and the fifth harmonic of the electron cyclotron frequency. Its bandwidth can extend beyond 100KHz and the frequency of the BUM peak is given by the empirical formula $f_{BUM} = 2f_o - nf_{ce}$ for $n = 3, 4, 5$. Note that the shape of BUM is highly variable ([11] Leyser) and the above empirical relation is a rough estimate.

- Broad symmetrical structure (BSS) — It was discovered in 1989 at Tromsϕ. It appears as a symmetrical structure which composes of two roughly equal sidebands (BSS^- and BSS^+), as shown in Figure 8(d). The frequencies of the BSS peaks are both deviated 15 to 30KHz from the pump. The BSS feature has the narrowest pump frequency range of existence among all the primary features and occurs for pump frequencies falling into a 40KHz interval near $3f_{ce}$, which is very similar to the DP feature. However, the coexistence of DP and BSS has never been observed ([21] Stubbe 1990). Note that the ionospheric condition in 1989 was remarkably different from that encountered in the previous experiments (1984-1988), due to a strongly enhanced solar activity level.

- Quenching of DM — The DM feature is quenched when the pump frequency is very close to nf_{ce}, $n = 3, 4, 5, 6, 7$. The quenching range gets smaller for increasing cyclotron harmonics. For the third harmonic case, the quenching range is estimated to be about 10KHz, while for the seventh harmonic, it was measured to be 200Hz ([13] Leyser 1992) in Russia. When the quenching occurs, the only features appearing in the SEE spectrum are the continuum and occasionally the DP.



(a) (b)

Figure 9: (a) SEE spectrum to demonstrate higher order DMs; (b) SEE spectrum to demonstrate $\frac{1}{2}DM$ and $\frac{1}{2}UM$ (From [20] Stubbe 1984).

The growth rate of the continuum is much higher than that of the DM and BUM. The rise time of the DM and BUM features is comparable to the growth time of small-scale striations. Also, the DM was occasionally observed strong during conditions of spread F ([11] Leyser).

Besides the above primary SEE features, there are secondary SEE features which their occurance is less often and less stable, and their observation is less repeatable. They include the second and third DMs (2DM and 3DM), the upshifted peak (UP), the second DP, the "misplaced" DM and UM which appear near the half frequency offset of the usual DM and UM ($\frac{1}{2}DM$ and $\frac{1}{2}UM$), and the split DM. Figure 9 displays some of the secondary SEE features ([20] Stubbe 1984).

Similar SEE phenomena, especially the continuum, the DP and the DM, are also observed at other locations in the northern hemisphere (e.g. Arecibo, Alaska, and in Russia). At Arecibo, the SEE phenomena was observed to be weaker than at Tromsϕ partly because the ERP in the Tromsϕ is about 2-3 times higher than in Arecibo ([23] Thide 1989). Another possible cause is the difference in magnetic dip between Tromsϕ and Arecibo. It is interesting, if possible, to have experimental data taken in the southern hemisphere for comparison.

In conclusion, the SEE spectral features are sensitive to small variations of the pump frequency around the electron cyclotron frequencies and to the ionospheric conditions. Figure 10 shows the temporal change of SEE spectrum measured at Arecibo. The pump frequency is slightly above the fifth harmonic of f_{ce}. Another temporal dependence of SEE can be seen from Figure 8(a) and (b), which were recorded in the same day but at different time.

The stimulated electromagnetic emissions have been proposed to use as a diagnostic tool to measure the maximum electric field strength in the reflection region of the pump wave and to measure the geomagnetic field where the upper hybrid frequency equals an electron cyclotron harmonic ([11]

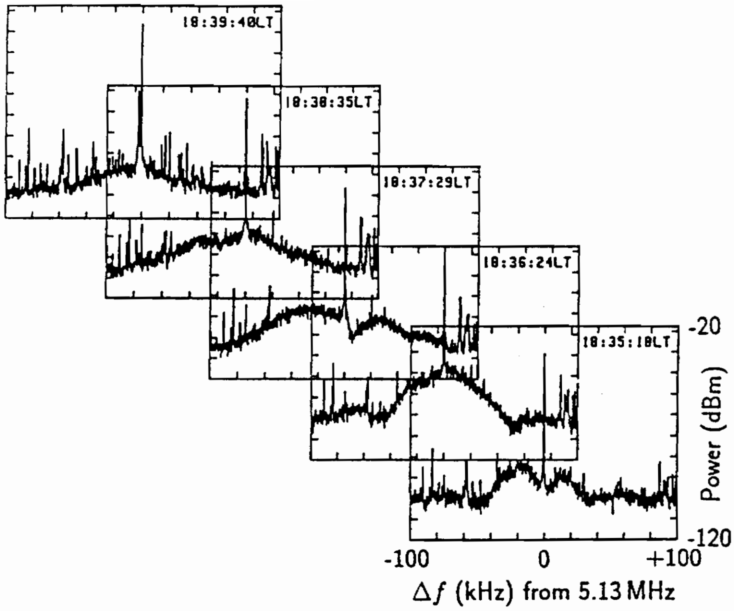


Figure 10: Five consecutive SEE spectra recorded at Arecibo (From [23] Thide 1989).

Leyser). Recently, Leyser conducted an experiment in Russia to determine the local magnetic field strength with an accuarcy of 1nT ([13] Leyser 1992).

3 Basic Plasma theory

A plasma is basically a system of N charged particles which are coupled to one another via their self-consistent electric and magnetic fields. In this chapter, we will review some of the basic plasma theory. The aim is to define and develop the important concepts that will appear in the discussion of stimulated electromagnetic emissions (SEE) and numerical simulation of SEE. The important concepts we discuss include physical plasma parameters such as plasma frequency, cyclotron radius and frequency, and Debye length. We then go on to discuss general wave propagation concepts in a plasma such as the dielectric tensor, O-wave and X-wave, cutoffs and resonances, hybrid frequencies, R-cutoff and L-cutoff. These basic ideas will then be used to described more complicated process that produce SEE. Due to space limitation, we cannot be rigorous in every detail. Fundamentals will be emphasized. But good references will be supplied whenever it is appropriate for more details.

3.1 Electrostatic electron plasma oscillation (Langmuir wave)

Electrostatic oscillations in a plasma were first discussed by Tonks and Langmuir in 1929. Here we discuss the high frequency electron oscillations which are too rapid for the heavy ions to follow. Thus the ions are treated as positive fixed charges.

Let us consider only two species, one is electrons and the other is singly charged positive ions. The density n_o of positive ions is uniform. Initially, the electrons also have uniform density n_o, but let us suppose that each electron is displaced in the x-direction by a small distance ξ which is independent of y- and z-coordinates and is zero on the plasma boundaries. The displacement of electrons disturbs the neutral plasma, producing a charge in each volumn element $\Delta x \Delta y \Delta z$:

$$\begin{aligned} \delta\rho\Delta x\Delta y\Delta z &= -n_o q_e \Delta y \Delta z \left[\xi - \left(\xi + \frac{\partial \xi}{\partial x}\Delta x\right)\right] \\ &= \Delta x \Delta y \Delta z n_o q_e \frac{\partial \xi}{\partial x} \end{aligned} \tag{7}$$

where q_e is the magnitude of an electron charge. The motion of the electrons produces an electric field E(x,t) which, because of the symmetry of the problem, is in x-direction. Thus using Possion's equation, we have

$$\frac{\partial E}{\partial x} = \frac{n_o q_e}{\epsilon_o}\frac{\partial \xi}{\partial x} \tag{8}$$

Then, integrating, we obtain

$$E = \frac{n_o q_e}{\epsilon_o}\xi \tag{9}$$

The force on each electron is $-q_e E$, which is proportional to the displacement ξ. It is also seen to be a restoring force. Thus, each electron oscillates about its original position with simple harmonic motion. The equation of motion for each electron is,

$$m_e \frac{d^2\xi}{dt^2} + \frac{n_o q_e^2}{\epsilon_o}\xi = 0 \tag{10}$$

The plasma frequency ω_p is defined, therefore, by

$$\omega_p = \sqrt{\frac{n_o q_e^2}{m_e \epsilon_o}} \tag{11}$$

where m_e is the electron mass. Substituting the numerical constants in the ionosphere, we get,

$$f_p = 8.966\sqrt{n_o} \tag{12}$$

where f_p and n_o are specified in Hz and m^{-3} respectively. A more thorough discussion of electron plasma oscillation can be found in many introductory texts such as [3] Chen.

3.2 Gyromotion of a charged particle in uniform magnetic field

The orbit of a particle of charge q moving in a prescribed electric and magnetic field may be calculated directly from the Lorentz force equation :

$$\mathbf{F} = \mathrm{q}(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{13}$$

First, let us consider a uniform magnetic field without any applied electric field. The Lorentz force is always at right angles, as dictated by the cross product in Lorentz force equation, to the velocity $\mathbf{v}$ of the charged particle. Hence its kinetic energy remains constant :

$$KE = \frac{1}{2} m_p v^2 = constant \tag{14}$$

where m_p is the mass of the particle. It is convenient to resolve the velocity $\mathbf{v}$ into two components: $v_{\parallel}$, parallel to $\mathbf{B}$, and $v_{\perp}$, in the plane perpendicular to $\mathbf{B}$. Since $v_{\parallel}$ is unaffected by the field, $KE_{\parallel} = \frac{1}{2} m_p v_{\parallel}^2$ remains constant too. It follows that $KE_{\perp} = \frac{1}{2} m_p v_{\perp}^2 = KE - KE_{\parallel}$ is also a constant of the motion. The Lorentz force provides a centripetal acceleration. Thus,

$$q v_{\perp} B = \frac{m_p v_{\perp}^2}{R} \tag{15}$$

The cyclotron radius of the orbit R (also called the Larmor radius) is given by,

$$R = \frac{m_p v_{\perp}}{|q| B} \tag{16}$$

and the cyclotron frequency Ω (also called the Larmor frequency) is,

$$\Omega = \frac{v_\perp}{R} = \frac{qB}{m_p}. \tag{17}$$

The cyclotron frequency carries the signs of the charge and the magnetic field. Thus opposite charges gyrate in opposite direction in the same magnetic field. The complete motion of the charged particle is described as a gyration of the particle in a circular orbit superimposed on the uniform motion of the orbit center, or guiding center, along a uniform magnetic field line, which results in a helical motion.

The magnetic field acts to confine the plasma by bending the particles in circular orbits. Of course no confinement is observed in the field direction. For ions and electrons of the same kinetic energy $KE_\perp$, the electrons gyrate in much smaller orbits, the ratio of the two Larmor radii being equal to the square root of the mass ratio. For parameters in the upper ionosphere, $B \approx 0.065 mT$ and the temperatures of electrons and ions (O^+) are 1000K and 250K respectively. Their corresponding cyclotron frequency and cyclotron radius are :

$$\Omega_e = -11.4\times10^6 \text{ rad/s } (f_{ce}=1.8\text{MHz}), \quad R_e=10.8 \text{ mm}$$

$$\Omega_i = 389 \text{ rad/s } (f_{ci}=61.8\text{Hz}), \quad R_i=0.92 \text{ m}.$$

where $f_c=|\Omega|/2\pi$ and $k_B T = m v_t^2$.

If an electric field and a magnetic field, both are uniform, are simultaneously applied to a plasma, and the electric field is perpendicular to the magnetic field, the case is more complicated than above. We will briefly mention the net effect. Serious readers may refer to, for examples, [3] Chen and [14] Reitz for details.

One may manipulate the Lorentz force equation to show that the total motion of the particle is made up of three terms : (a) constant velocity parallel to **B** , (b) gyration about the magnetic field lines, and (c) a constant drift velocity at right angles to both **E** and **B** . The last one is termed as E cross B drift.

If both electric and magnetic fields are nonuniform in space, the situation will become more complicated. In fact, a monotonically increasing magnetic field acts like a converging "lens" for the charged particles. That is the underlying principle of magnetic mirror which is often used to confine a plasma. A typical application is a fission reaction chamber. Again more details can be found in the two references cited in this section.

3.3 Debye shielding

One of the most important properties of a plasma is its tendency to remain electrically neutral. A slight imbalance in the space-charge densities gives rise to strong electrostatic forces which act, wherever possible, in the direction of restoring neutrality. On the other hand, if a plasma is deliberately subjected to an external electric field, the space-charge densities will adjust themselves so that the major part of the plasma is shielded from the field.

Let us consider a rather simple example. Suppose a spherical charge +Q is introduced into a plasma, thereby subjecting the plasma to an electric field. Actually, the charge +Q would be gradually neutralized because of being continuously struck by charged particles from the plasma, but if the charged object is physically very small, this will take an appreciable period of time. Meanwhile, electrons find it energetically favorable to move closer to the charge, whereas positive ions tend to move away. Under equilibrium conditions, the distribution of the charged particles against the potential energy is given by the Boltzmann distribution.

$$n_e = n_o \cdot e^{q_e(U_e - U_o)/k_B T_e} \quad (18)$$

$$n_i = n_o \cdot e^{-q_e(U_i - U_o)/k_B T_i} \quad (19)$$

where U is the local potential, U_o is the reference potential, and n_o is the electron density in regions where $U = U_o$. Here, we consider the ions are too heavy to move but form a uniform background of positive charge. So $U_i = U_o$. The potential U is obtained from the solution of Poisson's equation :

$$\frac{1}{r^2}\frac{d}{dr}(r^2\frac{dU}{dr}) = -\frac{1}{\epsilon_o}(n_i q_e - n_e q_e) = \frac{n_o q_e}{\epsilon_o}(e^{q_e(U-U_o)/k_B T} - 1). \quad (20)$$

This is a nonlinear differential equation and we cannot find its exact solution. So we consider an approximation which is rigorous at high temperature. If $k_B T \gg q_e U$, then $\exp(q_e U/k_B T) \approx 1 + q_e U/k_B T$, and

$$\frac{1}{r^2}\frac{d}{dr}(r^2\frac{dU}{dr}) = \frac{n_o q_e^2}{\epsilon_o k_B T}(U - U_o). \quad (21)$$

The solution is,

$$U = Uo + \frac{Q}{4\pi\epsilon_o r}e^{-r/\lambda_D}. \quad (22)$$

Here r is the distance from the spherical charge +Q, and λ_D is the Debye length which is given by,

$$\lambda_D = \sqrt{\frac{\epsilon_o k_B T}{n_o q_e^2}} \quad (23)$$

The redistribution of electrons in the gas is such as to screen out +Q completely in a distance of a few λ_D. Another useful formula which is used more often than the above one for λ_D is given by,

$$\omega_p \cdot \lambda_D = v_t \quad (24)$$

where v_t is the rms thermal velocity which is given by $k_B T = m v_t^2$. In the upper ionosphere, at 250km, the electron plasma frequency is about 7MHz and the electron temperature is about 1000K. The corresponding Debye length is about 2.8mm.

An ionized gas is called a plasma if it can satisfy the following three criteria :

- the Debye length is much smaller than other physical dimensions of interest, for example the plasma system length;
- there are a lot of charged particles inside a "Debye sphere" whose radius equals a Debye length, in order to enable the Debye shielding to be statistically valid; and
- the collisional frequency between charged and neutral particles is small compared with the frequency of typical plasma oscillations. Equivalently, $\omega_p \tau > 1$, where τ is the mean time between collisions with neutral atoms.

The upper ionosphere meets all these three criteria, therefore it is treated as a plasma.

3.4 Dielectric tensor and wave propagation in a cold magnetized plasma

We will now discuss the basics of electromagnetic wave propagation in a uniform magnetized plasma. The primary concern is the dispersion characteristic which relates propagation frequency to wave number. The cold plasma, that is the temperature is zero for all particle species, will be considered first because of its relatively simple derivation. After that, various approximation techniqes are introduced to deal with finite-temperature plasma.

The cold plasma dispersion was first published by Appleton in 1927 and 1932. Because Hartree influenced the publication of the 1932 derivation, although he added nothing to the result, it is sometimes called the Appleton-Hartree dispersion relation[1].

We start to derive the dielectric tensor of a cold plasma from the time-harmonic form of Maxwell's Equations. In this chapter, we always assume $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ variations to all continuous physical quantities. Then from the modified Ampere's law,

$$\nabla \times \mathbf{H} = \mathbf{J} - \mathrm{i}\omega\epsilon_o\mathbf{E} = -\mathrm{i}\omega\mathbf{D} \tag{25}$$

Here, the current density $\mathbf{J}$ is not a conduction current, but is a convection current generated by the motion of the charged particles. Then, the dielectric tensor $\mathbf{K}$ is defined by,

$$\mathbf{D} = \epsilon_o\mathbf{K}\cdot\mathbf{E} = \mathbf{E} + \frac{\mathrm{i}}{\omega}\mathbf{J} \tag{26}$$

As we will see later, we need a tensor description for the electric displacement and the electric field because the background magnetic field turns the uniform plasma into a anisotropic medium. The gyromotion and the interaction of the charged particles cause the anisotropy.

[1] Actually, the history is more complicated than the above-mentioned because there are many versions of the story claiming the originality of the dispersion relation.

Without loss of generality, let the background magnetic field $\mathbf{B_O}$ in the z–direction. The current density which goes into the definition of the dielectric tensor is given in terms of macroscopic particle densities and velocities.

$$\mathbf{J} = \sum_{\mathrm{k}} \mathrm{n_k q_k} \mathbf{v}_{\mathrm{k}} \tag{27}$$

The equation of motion for a single particle of type k is,

$$m_k \frac{d\mathbf{v}_{\mathrm{k}}}{\mathrm{dt}} = q_k(\mathbf{E} + \mathbf{v}_{\mathrm{k}} \times \mathbf{B}) = -\mathrm{i}\omega \mathrm{m_k} \mathbf{v}_{\mathrm{k}} \tag{28}$$

Time-harmonic variation has been assumed. In the cold-plasma model, each particle of the plasma oscillates about a fixed position in space under the influence of the electromagnetic field of the wave. There are no external electric field and drift velocity. All dissipative effects including collisions are neglected, We also assume small perturbation in the electric and magnetic fields. Thus, the background magnetic field dominates the effect on the particles. The contribution from charged particles is second order effect. Under this assumption, we are allowed to replace the total magnetic field in the Lorentz force equation by $\mathbf{B_O}$. Working out the cross product, we have,

$$\begin{aligned} v_x &= \frac{iq_k}{\omega m_k}(E_x + v_y B_o) \\ v_y &= \frac{iq_k}{\omega m_k}(E_y - v_x B_o) \\ v_z &= \frac{iq_k}{\omega m_k} E_z \end{aligned} \tag{29}$$

Note that only the first two velocities are coupled together. After substitution, the solution for the first two velocities is,

$$\begin{aligned} v_x &= \frac{iq_k}{\omega m_k}\left(E_x + i\frac{\Omega_k}{\omega}E_y\right)\left(1 - \frac{\Omega_k^2}{\omega^2}\right)^{-1} \\ v_y &= \frac{iq_k}{\omega m_k}\left(E_y - i\frac{\Omega_k}{\omega}E_x\right)\left(1 - \frac{\Omega_k^2}{\omega^2}\right)^{-1} \end{aligned} \tag{30}$$

where Ω_k is the cyclotron frequency for the type k species. Now we have the three velocities in terms of frequency. After substituting these three velocities into the current densities and then into the equation defining the dielectric cold tensor, we get,

$$\begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix} = \epsilon_o \begin{pmatrix} E_x(1 - \sum \frac{\omega_{pk}^2}{\omega^2 - \Omega_k^2}) - iE_y \sum \frac{\Omega_k}{\omega}\frac{\omega_{pk}^2}{\omega^2 - \Omega_k^2} \\ E_y(1 - \sum \frac{\omega_{pk}^2}{\omega^2 - \Omega_k^2}) + iE_x \sum \frac{\Omega_k}{\omega}\frac{\omega_{pk}^2}{\omega^2 - \Omega_k^2} \\ E_z(1 - \sum \frac{\omega_{pk}^2}{\omega^2}) \end{pmatrix} \tag{31}$$

where ω_{pk} is the plasma frequency for the type k species and the sums are over the species. Hence, the dielectric tensor is,

$$\mathbf{K} = \begin{pmatrix} 1-\sum \frac{\omega_{pk}^2}{\omega^2-\Omega_k^2} & -i\sum \frac{\Omega_k}{\omega}\frac{\omega_{pk}^2}{\omega^2-\Omega_k^2} & 0 \\ i\sum \frac{\Omega_k}{\omega}\frac{\omega_{pk}^2}{\omega^2-\Omega_k^2} & 1-\sum \frac{\omega_{pk}^2}{\omega^2-\Omega_k^2} & 0 \\ 0 & 0 & 1-\sum \frac{\omega_{pk}^2}{\omega^2} \end{pmatrix} \tag{32}$$

Sometimes, the entries of the dielectric tensor are written in short form.

$$\mathbf{K} = \begin{pmatrix} K_1 & K_2 & 0 \\ -K_2 & K_1 & 0 \\ 0 & 0 & K_3 \end{pmatrix} = \begin{pmatrix} S & -iD & 0 \\ iD & S & 0 \\ 0 & 0 & P \end{pmatrix} \tag{33}$$

where

$$S(=K_1) = 1-\sum_k \frac{\omega_{pk}^2}{\omega^2-\Omega_k^2} \tag{34}$$

$$D(=-iK_2) = \sum_k \frac{\Omega_k}{\omega}\frac{\omega_{pk}^2}{\omega^2-\Omega_k^2} \tag{35}$$

$$P(=K_3) = 1-\sum_k \frac{\omega_{pk}^2}{\omega^2} \tag{36}$$

The S,D,R,L and P notation are introduced by [16] Stix. The Stix symbols are mnemonics for sum, difference, right, left and plasma terms, respectively. Here S and D are defined again by the R and L terms as follows :

$$S = \frac{1}{2}(R+L) \tag{37}$$

$$D = \frac{1}{2}(R-L) \tag{38}$$

$$R = 1-\sum_k \frac{\omega_{pk}^2}{\omega^2}\frac{\omega}{\omega+\Omega_k} \tag{39}$$

$$L = 1-\sum_k \frac{\omega_{pk}^2}{\omega^2}\frac{\omega}{\omega-\Omega_k} \tag{40}$$

Here, we would like to mention the difference between a plane electromagnetic wave and an electrostatic wave. A plane wave has its propagation vector **k** perpendicular to the plane containing both the time-varying electric and magnetic fields. On the other hand, an electrostatic wave propagates with its wave vector parallel to the electric field. Thus, a plane wave is a transverse wave while an electrostatic plasma wave is a longitudinal wave.

Having obtained the dielectric tensor **K**, we can solve Maxwell's Equations for plane waves. We have,

$$\begin{aligned} i\mathbf{k}\times\mathbf{E} &= i\omega\mathbf{B} \\ i\mathbf{k}\times\mathbf{B} &= -i\omega\epsilon_o\mu_o\mathbf{K}\cdot\mathbf{E} \end{aligned} \tag{41}$$

It is convenient to introduce the dimensionless vector $\mathbf{n}$ which has the direction of the wave vector $\mathbf{k}$ and has the magnitude of the refractive index,

$$\mathbf{n} = \frac{\mathbf{k}c}{\omega} \tag{42}$$

Then, the two curl equations are simplified to,

$$\mathbf{n} \times (\mathbf{n} \times \mathbf{E}) + \mathbf{K} \cdot \mathbf{E} = 0 \tag{43}$$

Again, without loss of generality, we assume $\mathbf{n}$ lies on the x–z plane and intersects the background magnetic field $\mathbf{B_O}$ with an angle θ. Then we decomposite the propagation vector $\mathbf{n}$ into components $n_x = n\sin\theta$ and $n_z = n\cos\theta$, and use the vector identity,

$$\begin{aligned} \mathbf{n} \times (\mathbf{n} \times \mathbf{E}) &= \mathbf{n}(\mathbf{n} \cdot \mathbf{E}) - \mathbf{E}(\mathbf{n} \cdot \mathbf{n}) \\ &= n^2 \begin{pmatrix} (\sin^2\theta - 1)E_x + \sin\theta\cos\theta E_z \\ E_y \\ \sin\theta\cos\theta E_x + (\cos^2\theta - 1)E_z \end{pmatrix} \end{aligned} \tag{44}$$

Substituting into the wave equation, we get

$$\begin{pmatrix} S - n^2\cos^2\theta & -iD & n^2\cos\theta\sin\theta \\ iD & S - n^2 & 0 \\ n^2\sin\theta\cos\theta & 0 & P - n^2\sin^2\theta \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = 0 \tag{45}$$

In order to have a nontrivial solution, the determinant of coefficients must vanish. This condition gives the dispersion relation,

$$An^4 - Bn^2 - C = 0 \tag{46}$$

where

$$A = S\sin^2\theta + P\cos^2\theta \tag{47}$$

$$\begin{aligned} B &= PS(1 + \cos^2\theta) + S^2\sin^2\theta - D^2\sin^2\theta \\ &= RL\sin^2\theta + PS(1 + \cos^2\theta) \end{aligned} \tag{48}$$

$$C = (S^2 - D^2)P = PRL \tag{49}$$

The solution to the biquadratic eqaution is,

$$n^2 = \frac{B \pm F}{2A} \tag{50}$$

where

$$F^2 = B^2 - 4AC = (RL - PS)^2\sin^4\theta + 4P^2D^2\cos^2\theta \tag{51}$$

The dispersion relation was put into another form in terms of angle by Astrom and Allis ([16] Stix). Substitute the long expressions of A,B and C back into the biquadratic equation and then group terms with $\sin^2\theta$ and $\cos^2\theta$. After dividing, we have

$$\tan^2\theta = -\frac{P(n^2-R)(n^2-L)}{S(n^2-RL)(n^2-P)} \tag{52}$$

We will discuss some special cases of the general dispersion relation for a uniform magnetized plasma. Some of them are important in the discussion of later sections. First, we go back to the unmagnetized case. That is what the dispersion relation and the dielectric tensor should be for a plane wave propagating inside an unmagnetized plasma. The procedure is simple. We set the magnetic field B_O to zero and consider only one species, that is the electron. Furthermore, we assume the angle θ to be zero. So $A = P = S$, $B = 2S^2$, $C = S^3$, and $D = F = 0$. Here, the dielectric tensor consists only three non-zero and identical diagonal elements which is P. Therefore, without the magnetic field, the plasma is isotropic. The refractive index can be found to be,

$$n^2 = P = 1 - \frac{\omega_{pe}^2}{\omega^2} \tag{53}$$

We can put the above dispersion relation into another more familiar form by substituting $n = kc/\omega$.

$$c^2k^2 = \omega^2 - \omega_{pe}^2 \tag{54}$$

This equation tells us that the incident plane wave can propagate through the plasma only when the wave frequency is greater than the plasma frequency. Otherwise, the wave vector will be purely imaginary and reflection will occur at the point where the plasma frequency equals the wave frequency. This is a simplified account for the reflection of a HF radio wave from the ionosphere, which took a while for people to confirm that the cross Atlantic radio communication in 1920s was indeed due to the existence of the ionosphere. However, this picture is too simplified because the Earth has its background magnetic field and the ionosphere thus is a magnetized plasma.

In general, a magnetized plasma can support many types of waves. We summarize some special cases below.

1. Propagation parallel to $\mathbf{B_O}$ or $\theta = 0$. (The numerator of (52) must vanish.)
   (a) P=0 (Plasma oscillations)
   (b) $n^2 = R$ (wave with right-handed circular polarization)
   (c) $n^2 = L$ (wave with left-handed circular polarization)
2. Propagation perpendicular to $\mathbf{B_O}$ or $\theta = \pi/2$. (The denominator of (52) must vanish.)

(a) $n^2 = P$ (Ordinary wave or O-wave)

(b) $n^2 = RL/S$ (Extraordinary wave or X-wave)

Note that the dispersion of O-wave is the same as in an unmagnetized plasma discussed above because the electric field is parallel to the magnetic field. The wave polarizations in a plasma are with respect to the magnetic field direction, but not the wave vector direction. However, the problem of naming or differentiating the ordinary and extraordinary waves has created considerable confusion ([6] Hines). Usually, the names "ordinary" and "extraordinary" are adopted from the theory of optically birefringent crystals, but the basis on which these names are applied is far from uniform. In crystal optics, no ambiguity arises : the ordinary wave has the same directions of phase and group velocities, whereas the extraordinary wave does not. The two types of ionospheric radio waves, as a rule, have different directions for their group and phase velocities; hence the distinction used in optics is not readily applicable. Generally, the names are applied on the basis that the ordinary wave is the one less affected by the magnetic field, and hence resembles more closely the propagation of waves in a unmagnetized plasma. So one will expect the O-wave is reflected at the same height as it would be in the absence of magnetic field. For more physical insights about the above-mentioned wave modes, readers may refer to [3] Chen.

Inside a magnetized plasma, there exist many singularities which are classified into resonances and cutoffs. A resonance occurs when the refractive index goes to infinity or the wavelength goes to zero, while a cutoff is the condition when n goes to zero or the wavelength goes to infinity. Usually, a cutoff will result in reflection of the incident wave and a resonance will end up with absorption and/or reflection of the incident wave. The general condition for a resonance is given by (47) or (52) as,

$$A = 0 \quad \text{or} \quad \tan^2\theta = -\frac{P}{S} \tag{55}$$

and the general cutoff condition is given by (46) as,

$$C = PRL = 0. \tag{56}$$

Therefore, there are three cutoff conditons as seen from the above formula.

- $P = 0$ (Cutoff due to plasma oscillations)
- $R = 0$ (Right-handed cutoff)
- $L = 0$ (Left-handed cutoff)

As dictated by their names, the right-handed cutoff occurs only when the incident wave is right-handed polarized and the same argument can be applied to the left-handed cutoff.

For the resonances, we consider the following two special cases which are termed as principal resonances.

1. Parallel propagation ($\theta = 0$)

   Since P=0 is a cutoff, we require $S \to \infty$. But $S = (R+L)/2$, this can be satisfied for either $R \to \infty$ or $L \to \infty$. From the definitions of R and L (see (39) and (40)), only negative Ω_k can make $R \to \infty$. This implies that those charged particles are either electrons or negative ions. Generally, electrons will be a more important case and hence the right-handed resonance will be at the electron cyclotron frequency which is the highest cyclotron frequency among all the species. For the left-handed resonance, the above arguement will be reversed. When the wave frequency equals one of the ion cyclotron frequencies, we have a resonance. Here the ions are positive. Therefore for parallel propagation, the principal resonances are cyclotron resonances.

2. Perpendicular propagation ($\theta = \pi/2$)

   Since $P \to \infty$ is a trivial solution, we require $S \to 0$. These resonances are called hybrid resonances because they are generally involve some combination of Ω and ω_p. Here, we refer to X-wave resonance because an ideal O-wave does not have any resonance. Since the hybrid resonance is important for our later discussion, we will solve S=0 for a two-species plasma, one is electron and the other is positive ion.

$$S = 1 - \frac{\omega_{pe}^2}{\omega^2 - \Omega_e^2} - \frac{\omega_{pi}^2}{\omega^2 - \Omega_i^2} = 0 \tag{57}$$

$$\omega^2 = \frac{\omega_{pe}^2 + \Omega_e^2 + \omega_{pi}^2 + \Omega_i^2}{2} \pm \sqrt{\frac{\omega_{pe}^2 + \Omega_e^2 - \omega_{pi}^2 - \Omega_i^2}{2}^2 + \omega_{pe}^2\omega_{pi}^2} \tag{58}$$

$$\approx \frac{\omega_{pe}^2 + \Omega_e^2 + \omega_{pi}^2}{2} \pm \sqrt{\frac{\omega_{pe}^2 + \Omega_e^2 - \omega_{pi}^2}{2}^2 + \omega_{pe}^2\omega_{pi}^2} \tag{59}$$

   The approximation is justified for large mass ratio m_i/m_e and the fact that $\omega_p > \Omega$. The frequency with plus sign is called upper hybrid frequency, denoted by ω_{UH} and the one with minus sign is the lower hybrid frequency (ω_{LH}). They play crucial roles in the understanding of stimulated electromagnetic emissions which we will discuss in later section. A rough estimate, by assuming an overdense plasma[2], for the two hybrid resonance frequencies is,

$$\omega_{UH}^2 \approx \omega_{pe}^2 + \Omega_e^2 \tag{60}$$

$$\omega_{LH}^2 \approx \Omega_e\Omega_i \tag{61}$$

   The last approximation is obtained through Taylor series expansion of (58).

[2] That is a plasma with $\omega_{pe} \gg \Omega_e$.

It is a common practice to depict all these principal resonances and cutoffs in a Clemmow-Mullaly-Allis (CMA) diagram[3]. The x-axis is the variable $X=\omega_{pe}^2/\omega^2$ which depends on the density, and the y-axis is the variable $Y=|\Omega_e|/\omega$ which is a function of magnetic field. Then all cutoffs and resonances are plotted as CMA boundaries. For more details, one may refer to [16] Stix, [17] Swanson, or [3] Chen.

The analysis of cold plasma waves, although very complicated already, leaves out a great amount of physics which relates to finite temperature effects such as removal of singularities from the cold plasma dispersion relation. These effects may be included in varying degrees of approximation because exact treatment is impossible for every particle in the plasma. A well-known example is the thermal corrections for plasma oscillations proposed by Vlasov (1938) and Bohm and Gross (1949). In the next section, we will discuss some major techniques to deal with finite-temperature plasma.

In inhomogeneous plasma such as the ionosphere, the situation is much more complicated than the homogeneous one because the density variation can support more waves in the plasma. Moreover, the finite temperature and the background magnetic field, together with the inhomogeneity, can provide conditions for the interaction between waves and waves and interaction between waves and particles. But, for some not so complicated cases such as an unmagnetized plasma with slow varying inhomogeneity, the geometric optics or WKB method allows us to obtain good approximations to the exact solutions. In fact, it is a ray tracing method using the Eikonal equation. The criterion for valid approximation is that the wave number $k(\mathbf{r})$ is slowly varying. Otherwise, the geometric optics approximation breaks down.

Near a cutoff or resonance, the wavelength changes dramatically so that the geometric optics approximation is no longer applicable. Then we have to seek for different type of approximation or model which tailors for such a behavior. In general, the behavior near a cutoff is less complicated to handle and it is well known that the behavior is usually described by an Airy equation. This is referred as a full wave calculation ([2] Budden). The solutions to the Airy equation are Airy functions. Then, one can fix all the undetermined constants in the Airy function solutions by matching the asymptotic forms of solutions to the WKB solutions which model the regions other than cutoffs and resonances. Interested readers may refer to [16] Stix or [17] Swanson.

The analysis of resonances is intrinsically more difficult than the analysis of cutoffs because the physics of what resolves the resonance must be included in order to obtain physically meaningful results. In most cases, an isolated resonance results in absorption and/or reflection and it leads to the topic of mode conversion which will discuss in the section following the next.

[3] A more recent technique using three-dimensional plots for displaying all wave modes in a plasma is discussed in the PhD dissertation by M. André, Kiruna Geophysical Institute, University of Umeå, 1985.)

3.5 Kinetic and fluid descriptions of a plasma

In the last section, we derived the dispersion relation in a cold plasma by assuming the plasma is uniform and the thermal velocity is zero for all species. In other words, at t=0, we exactly know the position of each particle and their corresponding velocities which are zero. That is why we could easily write down the velocity components for each particle using Lorentz force equation in (29). However, if the temperature is finite, the velocities will spread out. Usually, it is expressed in terms of Maxwell-Boltzmann's distribution function. Here, we have N different initial velocities for N particles and we have to use N sets of Maxwell's Equations and Lorentz force equation to follow the subsequent motion of each particle. Their motions are coupled because the current density and the charge density in a particular set of Maxwell's Equations depend on the motions of other particles.

In the above formulation, we have a problem to assign the random velocity to each particle according to its velocity distribution function. So the concept of phase space is introduced to take care of distribution. In Liouville formulation, we treat the three positions and the three velocities of each particle as independent variables, with time being a parameter. Thus, in N-particle system, we have 6N+1 dimensions. Each particle has its own trajectory in its phase space. Liouville equation simply says that the distribution function in 6N+1 dimensions is conserved with respect to time. Details about the Liouville formulation can be found in [12] Nicholson. But, it is impossible for us to carry out all these exact calculations in a finite-temperature plasma. We never use Liouville equation in plasma calculation because it yields much more information than we want. We do not need to know the exact position and velocity of a particle at a particular time. Instead, we want to know some average and collective behaviors of the plasma, such as the overall drift velocity inside a plasma. These are statistical results. Therefore, the exact formulation is a valuable starting point to derive a reduced statistical description with appropriate approximations to yield practical information. A very good example is the Vlasov equation which is reduced to six phase space dimensions using BBGKY hierarchy ([12] Nicholson).

Before we talk about the Vlasov equation, we first spend some time to look at the single particle distribution function $f(\mathbf{x}, \mathbf{v}, t)$ which is the simplest approximation to the Liouville distribution function. Here we have seven independent variables, that is in Cartesian coordinates, (x, y, z), (v_x, v_y, v_z) and t. For a given time t, the number of particles in a small volume in phase space is $f(\mathbf{x}, \mathbf{v}, t)d^3x d^3v$. Thus $f(\mathbf{x}, \mathbf{v}, t)$ is the number density of particles in phase space. If we integrate $f(\mathbf{x}, \mathbf{v}, t)$ over the entire velocity space, we get the conventional number density $n(\mathbf{x}, t)$ defined by the number of particles in a given volume $\Delta x \Delta y \Delta z$.

$$\int f(\mathbf{x}, \mathbf{v}, \mathrm{t}) \mathrm{d}^3 \mathrm{v} = \mathrm{n}(\mathbf{x}, \mathrm{t}) \tag{62}$$

If we integrate $f(\mathbf{x}, \mathbf{v}, \mathrm{t})$ over the entire spatial volume, we obtain the familiar velocity distribution $f_v(\mathbf{v}, \mathrm{t})$ of the particles at time t.

$$\int f(\mathbf{x}, \mathbf{v}, \mathrm{t}) \mathrm{d}^3 \mathrm{x} = \mathrm{f_v}(\mathbf{v}, \mathrm{t}) \tag{63}$$

So if we go on to integrate over all the phase space, we will get,

$$\int \int f(\mathbf{x}, \mathbf{v}, \mathrm{t}) \mathrm{d}^3 \mathrm{x} \mathrm{d}^3 \mathrm{v} = \mathrm{N} \tag{64}$$

which is the total number of particles in the plasma. Sometimes the distribution function $f(\mathbf{x}, \mathbf{v}, \mathrm{t})$ is normalized by N so that the new distribution function is the probability density of finding particles in phase space.

3.5.1 Kinetic description (Vlasov Equation)

The simplest approximation in BBGKY hierarchy is the Boltzmann's equation. It is just the conservation of particles or probability density where the left hand side gives the rate of change following the trajectory of a particle in six-dimensional phase space and the right hand side represents the rate that trajectories are terminated through collisions with corresponding new trajectories started so that particles are conserved.

$$\frac{df_j(\mathbf{x}, \mathbf{v}, \mathrm{t})}{\mathrm{dt}} = \frac{\partial f_j}{\partial t} + \mathbf{v} \cdot \nabla_{\mathrm{x}} \mathrm{f_j} + \mathbf{a} \cdot \nabla_{\mathrm{v}} \mathrm{f_j} = \left(\frac{\mathrm{df_j}}{\mathrm{dt}} \right)_{\mathrm{c}} \tag{65}$$

The subscript j refers to the j^{th} species of particles and the two gradients are with respect to the spatial and velocity variables, indicated by their subscripts. The term **a** is the acceleration of particles. The Boltzmann's equation is originally derived for the kinetic theory of neutral gases. Hence the collisions are usually understood to be binary or head-on collisions. However, this is not the case in the plasma because we are dealing with long range force, i.e. the electromagnetism. The particle trajectory can be changed even without any in-touch collisions with other particles. In such a situation, we have to use random walk or Fokker-Planck formulation to find out the collisional term on the right side of Boltzmann's equation for Coulomb interactions ([17] Swanson or [12] Nicholson). Nevertheless, there are certain cases that the head-on collisions are important. When a gas is weakly ionized, the collisions between neutral molecules and charged particles are very often. The D-layer in the ionosphere is such an example.

In many plasmas, the collisional effect is not important. The cases we are considering in F region is one of the examples. This is obvious when we compare the collision frequency ($< 1KHz$) in Figure

7 with the plasma frequency ($\sim 5MHz$) at 200km. It was realized by Vlasov in 40s. Therefore, we can drop the collisional term in the Boltzmann's equation. The resulting equation is called the Vlasov equation.

$$\frac{df_j(\mathbf{x},\mathbf{v},\mathrm{t})}{\mathrm{dt}} = \frac{\partial f_j}{\partial t} + \mathbf{v}\cdot\nabla_{\mathrm{x}}\mathrm{f_j} + \frac{\mathrm{q_j}}{\mathrm{m_j}}(\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot\nabla_{\mathrm{v}}\mathrm{f_j} = 0 \tag{66}$$

for each species where $\dot{\mathbf{x}} = \mathbf{v}$ and $\dot{\mathbf{v}} = \mathbf{a}$. Here we substitute the Lorentz force equation into the acceleration term. Together with the Maxwell's Equations,

$$\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial\mathrm{t}} \tag{67}$$

$$\nabla\times\mathbf{H} = \mathbf{J} + \frac{\partial\mathbf{D}}{\partial\mathrm{t}} \tag{68}$$

$$\nabla\cdot\mathbf{D} = \rho \tag{69}$$

$$\nabla\cdot\mathbf{B} = 0 \tag{70}$$

$$\rho = \sum_j q_j \int f_j d^3v \tag{71}$$

$$\mathbf{J} = \sum_j \int \mathbf{v}\,\mathrm{f_j}\,\mathrm{d^3v} \tag{72}$$

we can deal with most of the analyses of plasmas if the collisional effect is not dominant. This type of fomulation is referred as kinetic treatment because we deal with distribution function. Compared with Liouville formulation, Vlasov formulation reduces dimensions from 6N+1 to seven by giving up details of the exact positions and velocities of all particles. The field quantities are macroscopic too.

Although the Vlasov-Maxwell equations are simplified version of the Liouville formulation, they are still nonlinear and in most cases, we have to use various approximations to obtain analytic solutions. One of the most often used techniques is linearization for studying small amplitude waves. This give us the linear Landau damping which is a collisionless damping effect for waves propagating in a plasma ([17] Swanson, [16] Stix, [3] Chen or [12] Nicholson).

3.5.2 Plasma as fluids

We can derive the fluid description of a plasma by taking moments of the Vlasov equation ([3] Chen, [12] Nicholson, or [17] Swanson). When we do that, we again throw away some details contained in the distribution function. Since $f\nabla_x\mathbf{v} = 0$ and $f\nabla_v\mathbf{a} = 0$ ($\mathbf{x}$ and $\mathbf{v}$ are independent variables), we can rewrite the Vlasov or collisionless Boltzmann equation as,

$$\frac{\partial f}{\partial t} + \nabla_x(\mathbf{v}\,\mathrm{f}) + \nabla_{\mathrm{v}}(\mathbf{a}\,\mathrm{f}) = 0 \tag{73}$$

Here we drop the subscript "j". Now we introduce a weighting function Q($\mathbf{v}$) and define the moment by an average over the velocity space as,

$$\langle Q(\mathbf{v})\rangle = \frac{\int \mathrm{Q\,f\,d^3v}}{\int \mathrm{f\,d^3v}} = \frac{1}{\mathrm{n}}\int \mathrm{Q\,f\,d^3v} \tag{74}$$

where n($\mathbf{r}$,t) is the number density mentioned above. We multiply the Vlasov equation by Q and integrate over velocity, we have

$$\int Q\frac{\partial f}{\partial t}d^3v + \int Q\nabla_x(\mathbf{v}\,\mathrm{f})\mathrm{d^3v} + \int \mathrm{Q}\nabla_{\mathrm{v}}(\mathbf{a}\,\mathrm{f})\mathrm{d^3v} = 0 \tag{75}$$

Since Q is a function of $\mathbf{v}$ only, this becomes,

$$\frac{\partial}{\partial t}\int Qfd^3v + \nabla_x\int Q\,\mathbf{v}\,\mathrm{f\,d^3v} + \int \mathrm{Q}\nabla_{\mathrm{v}}(\mathbf{a}\,\mathrm{f})\mathrm{d^3v} = 0 \tag{76}$$

The first term is simply $\partial(n\langle Q\rangle)/\partial t$ and the second is $\nabla_x(n\langle Q\mathbf{v}\rangle)$. Using one of the vector identities, the third term expands to,

$$\begin{aligned}\int Q\nabla_v(\mathbf{a}\,\mathrm{f})\mathrm{d^3v} &= \int[\nabla_v(Q\,\mathbf{a}\,\mathrm{f}) - \mathrm{f}\,\mathbf{a}\cdot\nabla_{\mathrm{v}}\mathrm{Q}]\mathrm{d^3v} \\ &= \oint_{S_v} Q\,f\,\mathbf{a}\cdot\mathrm{d}\mathbf{S}_{\mathrm{v}} - \int \mathrm{f}\,\mathbf{a}\cdot\nabla_{\mathrm{v}}\mathrm{Q\,d^3v}\end{aligned} \tag{77}$$

The surface integral in velocity space vanishes because we assume the distribution vanishes for $v \longrightarrow \infty$. We can then write (76) as,

$$\frac{\partial}{\partial t}(n\langle Q\rangle) + \nabla_x(n\langle Q\mathbf{v}\rangle) - \mathrm{n}\langle\mathbf{a}\cdot\nabla_{\mathrm{v}}\mathrm{Q}\rangle = 0 \tag{78}$$

We first take the zeroth moment of the Vlasov equation by letting Q=1. Then $\langle Q\rangle = 1$ and $\langle Q\mathbf{v}\rangle = \langle\mathbf{v}\rangle = \mathbf{u}$, where $\mathbf{u}$ is the mean or average velocity of the fluid element. Then (78) leads to,

$$\frac{\partial n}{\partial t} + \nabla(n\mathbf{u}) = 0 \tag{79}$$

This is just the continuity equation. The subscript "x" of the del operator is dropped because in the following, it only refers to differentiation defined in the three spatial dimensions. For the first moment, we let $Q = mv_x$. Then $\langle Q\rangle = mu_x$ and $\nabla_v Q = mv_x\hat{\mathbf{v}}_x$, where $\hat{\mathbf{v}}_x$ is the unit vector in v_x-direction. So $\langle\mathbf{a}\cdot\nabla_{\mathrm{v}}\mathrm{Q}\rangle = \mathrm{m}\langle\mathrm{a_x}\rangle$. This leads to

$$\frac{\partial}{\partial t}(nmu_x) + \nabla(nm\langle\mathbf{v}\mathrm{v_x}\rangle) - \mathrm{nm}\langle\mathrm{a_x}\rangle = 0 \tag{80}$$

Now we let $\mathbf{v} = \mathbf{u} + \mathbf{w}$, where $\mathbf{w}$ measures the perturbation from the average velocity (e.g. thermal agitation) and $\langle\mathbf{w}\rangle = 0$. Then,

$$\langle\mathbf{v}\mathrm{v_x}\rangle = \langle(\mathbf{u}+\mathbf{w})(\mathrm{u_x}+\mathrm{w_x})\rangle = \mathbf{u}\mathrm{u_x} + \langle\mathbf{w}\mathrm{w_x}\rangle \tag{81}$$

The last term is

$$nm\langle a_x\rangle = q\int(\mathbf{E}+\mathbf{v}\times\mathbf{B})_{\mathrm{x}}\,\mathrm{f}\,\mathrm{d}^3\mathrm{v} = \mathrm{nq}(\mathbf{E}+\mathbf{u}\times\mathbf{B})_{\mathrm{x}} \tag{82}$$

By taking all three such component equations by letting $Q = mv_x, mv_y, mv_z$, we obtain the first moment equation,

$$\frac{\partial}{\partial t}(nm\mathbf{u}) + \nabla(\mathrm{nm}\mathbf{u}\cdot\mathbf{u}\mathrm{v_x}) + \nabla(\mathrm{nm}\langle\mathbf{w}\cdot\mathbf{w}\rangle) - \mathrm{nq}(\mathbf{E}+\mathbf{u}\times\mathbf{B}) = 0 \tag{83}$$

The term $nm\langle\mathbf{ww}\rangle$ is often called the stress tensor, denoted by $\mathbf{P}$ whose components $P_{ij} = nm\langle w_i w_j\rangle$ specify both the direction of motion and the component of momentum involved. It is a generalization of the scalar pressure to the anistropic pressure. Using the following identity and the continuity equation,

$$\nabla(nm\mathbf{u}\cdot\mathbf{u}) = \mathrm{nm}(\mathbf{u}\cdot\nabla)\mathbf{u} + \mathrm{m}\mathbf{u}\nabla\cdot(\mathrm{n}\mathbf{u}) \tag{84}$$

the above equation is generally written as,

$$nm\left[\frac{\partial\mathbf{u}}{\partial \mathrm{t}} + (\mathbf{u}\cdot\nabla)\mathbf{u}\right] + \nabla\cdot\mathbf{P} - \mathrm{n\,q}\,(\mathbf{E}+\mathbf{u}\times\mathbf{B}) = 0 \tag{85}$$

This equation is the fluid force equation or is sometimes called the momentum equation.

The momentum equation does not include any collisional effects. We can easily include neutral collisions by adding a simple term $nm(\mathbf{u}-\mathbf{u}_{\mathrm{o}})/\tau$ to (85), where $\mathbf{u}_o$ is the velocity of the neutral fluid and τ is the mean time between collisions.

(79) gives us the evolution of n as a function of $\mathbf{u}$, while (85) gives us the evolution of $\mathbf{u}$ as a function of $\mathbf{P}$. To find the evolution of $\mathbf{P}$, one needs the next higher moment and the process goes on. The infinity process is usually truncated after the first moment. The equation of state of a gas is used for the approximation of the pressure term. For the 1-dimensional case, the pressure is isotropic. From the theory of thermodynamics, we have equation of state,

$$P = \gamma n k_B T \tag{86}$$

where γ is the ratio of specific heats. For 1-D case, γ equals one for isothermal compression and equals 3 for adiabatic compression.

Now we can summarize the many-fluid description of plasma as follows. For each species, we have the continuity and momentum equations.

$$\frac{\partial n_j}{\partial t} + \nabla(n_j\mathbf{u}_j) = 0 \tag{87}$$

$$n_j m_j\left[\frac{\partial\mathbf{u}_j}{\partial t} + (\mathbf{u}_j\cdot\nabla)\mathbf{u}_j\right] + \nabla\cdot\mathbf{P}_j - n_j q_j(\mathbf{E}+\mathbf{u}_{\mathrm{j}}\times\mathbf{B}) = 0 \tag{88}$$

The pressure of each charged fluid is related to its density by an equation of state, which depends on the characteristic frequency and wavenumber of the process being considered. When $\omega/k \ll v_{tj}$, where v_{tj} is the thermal velocity of the fluid of the j^{th} type, the isothermal equation of state is valid. When $\omega/k \gg v_{tj}$, the adiabatic process should be used. When $\omega/k \sim v_{tj}$, the details of the velocity distribution of the charged particles are important. The fluid model is inadequate and we must return to the Vlasov equation.

Maxwell's Equations are related to the continuity and momentum equations through the charge and current densities. For our many-fluid model,

$$\rho = \sum_j n_j q_j \tag{89}$$

$$\mathbf{J} = \sum_j \mathrm{n_j q_j} \mathbf{u}_\mathrm{j} \tag{90}$$

This completes the formulation of the many-fluid model for plasmas.

For a plasma composed of electrons and one species of ions, the many-fluid model reduces to the well-known two-fluid model. If we go back to rederive the dispersion for the high-frequency electrostatic electron plasma oscillation using the fluid treatment, we will get ([3] Chen, [17] Swanson, or [10] Kruer),

$$\omega^2 = \omega_p^2 + 3k^2 v_{te}^2 \tag{91}$$

In dense ionized gases or conducting liquids, the collision frequency is sufficiently high so that the conduction of current in the plasma nearly obeys Ohm's law. Although there exist a large number of high-frequency jitter in the particle motion, the electrons and ions move in such a way that there is no separation of charge on the average. Then the mechanical motion of the system can be descibed in terms of a single conducting fluid with the usual hydrodynamic variables of density, velocity and pressure. We can combine the two fluids into one by,

$$\begin{aligned} \text{Mass density} : \rho_M(\mathbf{x}) &= m_e n_e(\mathbf{x}) + \mathrm{m_i n_i}(\mathbf{x}) \approx \mathrm{m_i n_i}(\mathbf{x}) && (92) \\ \text{Fluid velocity of center of mass} : \mathbf{v} &= \frac{m_i n_i \mathbf{v}_i + m_e n_e \mathbf{v}_e}{\rho_M} && (93) \\ \text{Total pressure} : P &= P_i + P_e && (94) \end{aligned}$$

This approximation describes the magnetohydrodynamics (MHD) formulation. However, we need to include the terms containing collisional effect in MHD.

We have reviewed most of the techniques which we deal with plasma. But, our success in obtaining complete analytical solutions is still very limited because most of the equations are nonlinear and involve more than one independent variables. Therefore, we cannot always solve the problem directly, especially when we deal with magnetized and inhomogeneous plasma. We have to separate the whole problem into parts. In other words, we create models according to physical processes. An example is the way we deal with the electrostatic electron plasma oscillations. We first assumed the heavy ions do not move. Then we concentrate on calculating the electron motion. Under the context of modelling by processes, we have to guess which processes are important to the phenomena we observed. Then we apply our analytic techniques to solve the formulated problems for each process. Hopefully we can get consistent results when we concatenate all the processes.

In addition to the analytical solutions, we can tackle the problems by numerical simulations. Here we simulate the plasma with one of the formulation we talked above, including following the trajectories of each particle. Then we can get a comprehensive picture, including the nonlinear effects, of the plasma, although computer simulation is limited by memory size, computational speed, nummerical errors and numbers of independent variables. We will discuss the numerical simulation in more details in chapter 4.

3.6 Mode conversion

Mode conversion is a phenomenon in which a wave of one type is linearly coupled into a wave of another type. In 1960s, many people set up experiments to heat the ionosphere. Soon they discovered that there were certain propagation angles for which some amount of the incident wave energy was absorbed inside the plasma. Later, as the diagnostic equipment advanced, people observed large cavity formations near the reflection point in the ionosphere (e.g. [25] Wong 1987), which may be caused by mode conversion. Today, mode conversion is regarded as a standard process of converting an electromagnetic wave into an electrostatic wave in an inhomogeneous plasma. We refer this as forward conversion. Of course, there is backward conversion, where an electrostatic plasma wave is converted into a radiating electromagnetic wave. Both of these two conversion processes are present in the SEE experiments. However, we isolate and consider the forward conversion because it is one possible mechanism that generates the SEE spectrum.

3.6.1 Linear mode conversion

In a homogeneous plasma, linear waves are not coupled, and propagate independently. In an inhomogeneous plasma, density variations and hence various singularities exist. When the incident wave perturbs the plasma, there are points, especially near resonances, at which two waves have the same wavenumber and matching phase velocity. The plasma will excite both waves and transfer some energy between the waves. This leads to the phenomenon of mode conversion where a wave of one type is linearly coupled into a wave of another type. Hence the mode conversion is sometimes referred as the linear conversion. The conversion efficiency can be 100%, though rarely.

From our previous discussion of resonances and cutoffs, a resonance occurs when the refractive index goes to infinity or the wavelength drops to zero. Then the propagation velocity also drops to zero. One may expect that a mode-converted plasma wave cannot propagate very fast and in general, it is an electrostatic wave which has a very slow propagation speed. However, if a cutoff is spatially close to a resonance, the conversion process will be more complicated. That is the case in the ionosphere.

Recall the dispersion relation $n^2 = F(\omega)$ obtained from the dielectric tensor analysis in the previous section. Since many ionospheric wave propagation researchers use X($= \omega_{pe}^2/\omega^2$) and Y($=|\Omega_e|/\omega$) variables to describe dispersion relation, we will modify the Stix notations such as S,D and P in order to explain Figure 11. Here, we deal with high frequency waves, so only electrons are considered. After substitutions, the five Stix notations are tranformed to,

$$S = 1 - \frac{X}{1-Y^2} \tag{95}$$

$$D = -\frac{XY}{1-Y^2} \tag{96}$$

$$P = 1 - X \tag{97}$$

$$R = 1 - \frac{X}{1-Y} \tag{98}$$

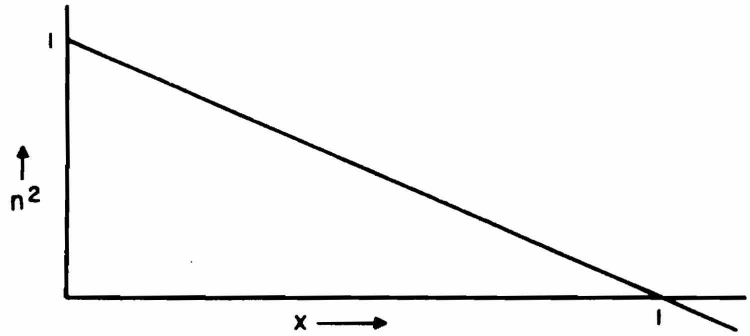
$$L = 1 - \frac{X}{1+Y} \tag{99}$$

The general dispersion relation for an arbitrary propagation angle is given by (50). But we want to demonstrate the simplest case below. For an ideal O-mode propagation, the newly transformed dispersion relation is,

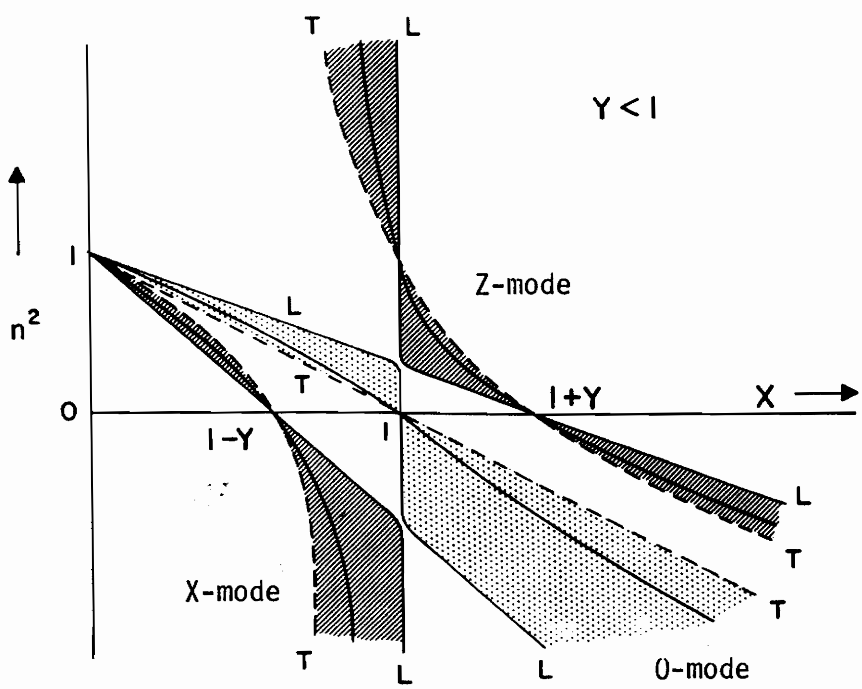
$$n^2 = 1 - X \tag{100}$$

and it is plotted in Figure 11(a). Obviously it is a straight line. For the other wave modes, the dispersion is depicted in Figure 11(b). The extreme cases (e.g. pure X-wave, R-wave, L-wave, etc.) are labeled with L and T, which stand for longitudinal (parallel) and transverse (perpendicular) propagations respectively. Also note that Y is less than one. This is typical in the ionosphere. All the three zeros (at $X = 1-Y$, $X = 1$, and $X = 1+Y$) of n^2 are cutoffs. They are invariant against propagation angle. There is only one resonance for each propagation angle. It corresponds to the approximate value of the upper hybrid frequency, where the mode conversion actually takes place. The resonance always lies between $X = 1 - Y^2$ and $X = 1$ because from (57) and (95). The order of these cutoff and resonance frequencies is : $\omega_L < \omega_{pe} < \omega_{UH} < \omega_R$, where ω_L and ω_R are the left-handed and right-handed cutoff frequencies respectively. There are no propagating wave modes in $n^2 < 0$ region, but they can be evanescent. The wave mode on the right side in Figure 11(b) is sometimes call the Z-mode or slow extraordinary mode (e.g. [16] Mjolhus 1990).

In the ionosphere, due to the existence of density gradient, right-handed circularly polarized wave



(a)



(b)

Figure 11: (a) Dispersion of an unmagnetized plasma. (b) Dispersion of a magnetized plasma with $Y < 1$. The solid and dashed lines indicate the limiting cases of longitudinal (L) and transverse (T) propagation respectively. The shades areas are the regions where the curves for other propagation angles lie. (Both from [6] Hines)

will be reflected at a point where $\omega = \omega_R$. This point is at a lower altitude than the plasma cutoff ($\omega = \omega_{pe}$) because for a fixed wave frequency, the right-handed cutoff occurs with a smaller electron density than the plasma cutoff. Thus only the left-handed circularly polarized wave can be reflected from the plasma cutoff point. But, no wave can be reflected from the left-handed cutoff point because we assume the density gradient is increasing with height. So from what we discussed in section 3.4 about the difference between O-wave and X-wave, O-wave is left-handed polarized and X-wave is right-handed polarized because the reflection height of the O-wave is apparently unaffected by the presence of magnetic field. The wave polarization is with respect to the magnetic field direction. It needs no change in the southern hemisphere where the magnetic field points upward, when the wave polarization refers to its wave vector. But, it is reversed in the northern hemisphere if the polarization is with respect to the wave propagation direction.

We are going to use Figure 11(b) to understand how mode conversion occurs in the ionosphere. Mode conversion involving X-wave has been shown to be inefficient in ionospheric heating experiments because the density gradient of the ionosphere is relatively small ([16] Mjolhus 1990) We will consider O-wave only because by its definition, only O-wave has access to the region of high-frequency plasma reflection (X=1). The mode conversion can be described as a two-stage process ([15] Mjolhus 1984 and [16] Mjolhus 1990). First, the incident O-wave is converted into a Z-mode at X=1 by tunneling — an effect occurs near a pair of resonance and cutoff when the incident "fast" wave is transformed into a "slow" wave ([16] Stix). The incident wave is reflected at X=1. However, no resonance occurs on the dispersion curve of the O-mode. Instead the resonance occurs on the Z-mode which can only exist inside the plasma. But, the O-mode can couple to the Z-mode, especially when the propagation angle θ is near the so-called critical angles[4]. There are two distinct critical angles that can result in complete coupling of O-mode into Z-mode. Their corresponding ray trajectories are depicted in Figure 12. It is computed by ray tracing method and the mode conversion is assumed to take place in northern hemisphere. The mode conversion that occurs with the ray on the right side is called southward process (positive critical angle) and the other one is northward process (negative critical angle). We consider only the southward process in this section. When the magnetic field and the wave vector are exactly aligned (purely parallel propagation), the reflection point is at $X = 1+Y$ because the dispersion curves of O-mode and Z-mode touch at $X = 1$.

[4] It is the angle of incidence of an electromagnetic wave with respect to the geomagnetic field, at which the coupling between the O-mode and Z-mode is 100%. The formula for the critical angle is given in [16] Mjolhus 1990 as $\sin\theta_c = \pm\sqrt{Y/(1+Y)}\sin\alpha$ for $0 \le \alpha \le \pi/2$. Note that the coordinate system used in Mjolhus' paper is different from that used in section 3.4. The density gradient is assumed in z-direction. The geomagnetic field intersects the z-axis with an angle α. The wave vector and the geomagnetic field are assumed in the same plane. Here θ is the angle between the wave vector and z-axis. The cone formed by the two critical angles is called the radio window.

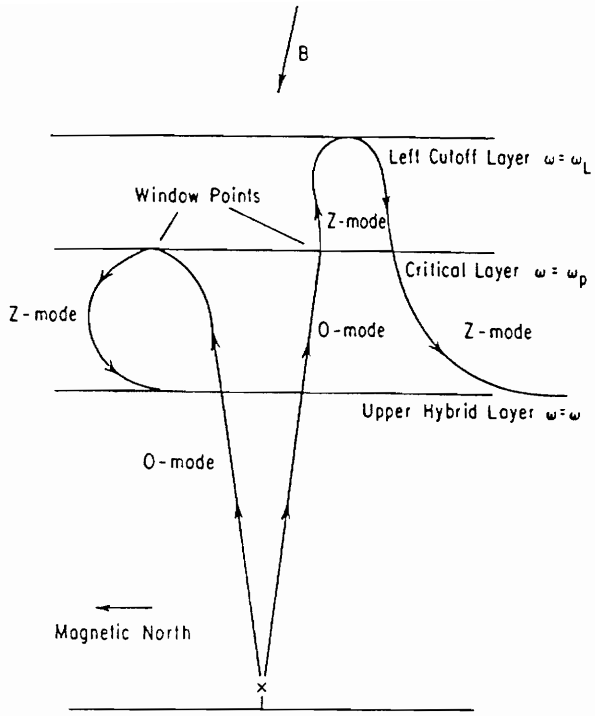


Figure 12: Ray trajectories for the illustration of mode conversion at critical angles (From [5] Close 1990).

When the magnetic field crosses the wave vector at a small angle θ, the incident O-wave is partially reflected at $X = 1$ and partially transmitted to Z-mode. A formula for the transmission coefficient is also given in [16] Mjolhus 1990. However, as the transmitted electromagnetic wave passes through $X = 1$, the wave mode changes from O-mode to Z-mode and the refractive index jumps from a low value to a high value. The wave vector is still directed upward. But the transmitted ray is bent at $X = 1$, as shown in Figure 12.

The second stage begins with the Z-mode wave propagating towards the left-handed cutoff ($X = 1+Y$). After reflection from $X = 1+Y$, the Z-mode wave propagates towards decreasing X. When it approaches the upper-hybrid resonance layer, it becomes gradually electrostatically polarized and the wave velocity[5] slows down. Thus it becomes 100% converted into an electrostatic wave. As seen from Figure 12, the mode-converted electrostatic wave (generally called upper hybrid wave) propagates parallel to the density stratification. All the mathematical derivations of the above process can be found in [15] Mjolhus 1984.

[5] All wave velocities mentioned in this chapter refer to the group velocities which possess relevant physical meaning for plasma waves.

So far, we use the result from cold plasma theory to qualitatively investigate the mode conversion. But the ionosphere is a finite-temperature plasma. We would not anticipate some singularity-type phenomena occur. Instead going to infinity at resonance, the refractive index will remain finite, though very large, in the ionosphere. Furthermore, the cutoffs and resonance will not be a single point. They are spread out over a finite region due to thermal effect. The full treatment of such a mode conversion in a warm plasma is more complicated. It involves modelling of the behaviors near a cutoff and a resonance, as well as the WKB method. Interested readers may refer to [16] Stix, [17] Swanson, and [15] Mjolhus 1984.

There are some numerical simulations illustrating the mode conversion process in an inhomogeneous magnetized plasma. They are [14] Lin 1982 and [5] Close 1990. In the former paper, linear density gradient is used, whereas in the latter papers, the authors assume a constant density with inhomogeneous background magnetic field. In both papers, the pump frequency is set at the second harmonics of the electron cyclotron frequency and X-wave is used. In order to see the mode conversion effect, they use 1-D electromagnetic simulation code. In both simulations, electrostatic plasma wave is generated around the upper hybrid point along the inhomogeneity. It shows the crucial role of the inhomogeneity in the mode conversion. Also in the simulations, we can see the reflection and absorption of the incident electromagnetic wave.

3.6.2 Direct conversion

Besides the mode conversion we discussed above, there is another possibility for the conversion of an electromagnetic wave into an electrostatic wave. It is termed as "direct conversion" when it was reported in [24] Wong 1981. The crucial part in this direct conversion is the pre-existing short-scale ($\sim 1m$) field-aligned irregularities which play the role of an in-situ converter, but not the density gradient. The existence and dynamics of the ionospheric irregularities have been well documented (e.g. [9] Kelley). Irregularities of a variety of scale sizes (varying over many orders of magnitude) exist in the different regions of the ionosphere. They arise from a number of sources such as wind, gravity, gradients, etc. The physical process of the direct conversion is as follows ([1] Antani 1991). The incident O-wave induces oscillatory electron drift in the upper-hybrid resonance zone. This induced electron velocity beats with the pre-existing density irregularities to generate a source current that plays the role of an in-situ antenna radiating the excited upper-hybrid wave. This process occurs whenever the following frequency and wave vector matching conditions are satisfied :

$$\omega_o = \omega_{UH} + \omega_n \tag{101}$$

$$\mathbf{k}_{\mathrm{o}} = \mathbf{k}_{\mathrm{UH}} + \mathbf{k}_{\mathrm{n}} \tag{102}$$

where the subscripts "o" and "n" refer to the pump and the irregularities, respectively. The matching conditions are same as in the parametric decay process, which will discuss in the following section. However, it differs from the usual parametric process in three subtle respects :

- it has no threshold power requirement;
- it is basically linear and no feedback is involved; and
- it leads to initial secular growth that is linear in time.

It can be distinguished from the conventional linear mode conversion process in such a way that it occurs independent of any standard plasma resonances. Also, since the direct conversion relies on these irregularities in the ionosphere, which vary with time and position, we can expect the direct conversion process is not a static process, but a process we will see a high degree of fluctuations. It may be correlated to some experimental observations of anomalous absorption of the incident electromagnetic wave in the ionospheric heating experiments (e.g. [8] Jones 1990).

Now, we are not in a position to decide which one of these two conversion processes is dominantly responsible for the generation of electrostatic wave in the upper-hybrid layer. For the conventional mode conversion, we can use electromagnetic simulation code to investigate some of the features. But for the direct conversion, we have to rely on measurement to understand some details about the process. In fact, we may be interested in the direction of the upper-hybrid wave generated because together with the information about the density irregularities, we may be able to decide which process is more likely. Here we do not exclude any other possibilities that may cause an electromagnetic wave converted into an electrostatic wave in the ionosphere.

For the back conversion process, considerably less literatures can be found. Since this is not the most important theory we count on, we will briefly mention the proposed theory in [17] Oya 1971 in the following. In a slightly inhomogeneous plasma, the Bernstein-mode[6] electrostatic wave can escape by being converted into the O-mode electromagnetic wave. Two reflections take place during this escape process.

3.7 Pondermotive force and parametric instability

Parametric instability is an important physical process to understand the SEE. Its generation is related to the interaction between the pondermotive force and in most cases, the ion density fluctuations. In the section, we will discuss the pondermotive force first. Then a simple case is used

[6] It will be briefly mentioned in section 3.8.

to demonstrate how the coupling between the pondermotive force and the ion density fluctuations causes parametric instability.

The pondermotive force

We begin with the pondermotive force or Miller effect on a single particle, which occurs in spatially varying high frequency electric fields, with or without an accompanying magnetic field ([12] Nicholson). Consider a charged particle oscillating in a high frequency electric field $E(t) = E_o \cos(\omega t)$. The motion is then a sinusoidal variation of distance with time. Now suppose the electric field has an amplitude that varies smoothly in space, $E(x,t) = E_o(x)\cos(\omega t)$, being stronger to the right and weaker to the left. Then the first oscillation brings the particle into regions of strong field, where it can be given a strong push to the left. When the field turns around, the particle is in a region of weaker field, and the push to the right is not as strong. The net result is a displacement to the left, which continues in succeeding cycles as an acceleration away from the region of strong field.

Mathematically, the Lorentz force equation is,

$$m\frac{d^2x}{dt^2} = qE_o(x)\cos(\omega t) \tag{103}$$

It is convenient to decompose x into a slowly varying component x_o, called the oscillation center and a rapidly varying component x_1, $x = x_o + x_1$. Here, x_o is a time average of the position x over the short time $2\pi/\omega$. Make a Taylor expansion of $E_o(x)$ about the oscillation center x_o, we have,

$$m(\ddot{x}_o + \ddot{x}_1) = q\left[E_o(x_o) + x_1\left(\frac{dE_o}{dx}\right)_{x_o}\right]\cos(\omega t) \tag{104}$$

Averaging (104) over time, we get,

$$m\ddot{x}_o = q\left(\frac{dE_o}{dx}\right)_{x_o}\langle x_1\cos(\omega t)\rangle_t \tag{105}$$

since $\langle E_o(x_o)\cos(\omega t)\rangle_t$ and $\langle x_1\rangle_t$ are zero. To obtain an equation for x_1, we note that $\ddot{x}_1 \gg \ddot{x}_o$ because x_1 is high frequency; moreover, in the spirit of the Taylor expansion, we have $E_o \gg x_1(dE_o/dx)$; therefore,

$$m\ddot{x}_1 = qE_o\cos(\omega t) \tag{106}$$

The solution is $x_1 = -(qE_o/m\omega^2)\cos(\omega t)$. Insert this in (105) and carry out the time average, we get,

$$\ddot{x}_o = -\frac{q^2E_o}{2m^2\omega^2}\left(\frac{dE_o}{dx}\right)_{x_o} \tag{107}$$

and so the pondermotive force $F_p = m\ddot{x}_o$ is,

$$F_p = -\frac{q^2}{4m\omega^2}\frac{d}{dx}(E_o^2) \tag{108}$$

Another interesting way to interpret the pondermotive force is that if we introduce the jitter speed $\tilde{v} = (\dot{x}_1)_{max} = qE_o/m\omega$; then

$$F_p = -\frac{m}{4}\frac{d}{dx}(\tilde{v}^2) \tag{109}$$

Notice the overall mass dependence in (108), so the pondermotive force acts much more strongly on electrons than on ions.

The above derivation is for a single particle motion. But, same principle can apply to a 3-D plasma using the continuity equation ([10] Kruer). The general formula for the pondermotive force is,

$$\mathbf{F}_p = -\frac{q_e^2}{4m_e\omega^2}\nabla|E_o(\mathbf{x})|^2 \tag{110}$$

A more complete derivation of the pondermotive force, including the magnetic field, can be found in [15] Schmidt.

The ionosphere is a magnetized plasma. It is impossible for us to perform a full treatment of the parametric instability here due to its complexity. Instead we will use a simple case, namely a one-dimensional uniform unmagnetic plasma, to illustrate the concept.

Coupling via ion density fluctuations

In the following, we consider the coupling of an electromagnetic wave into an electron plasma wave via ion density fluctuations. Let the electromagnetic wave be a spatially homogeneous oscillating electric field $\mathbf{E}_1 = \hat{x}E_1 exp(-i\omega t)$. In other words, the wave number of the electromagnetic wave is neglected on the assumption that it is much less than k, the wave number of the fluctuation in ion density. Since the frequency of an ion wave is much less than the frequency of an electromagnetic wave, we describe the ion density fluctuation as static modulation in the plasma density, $n = n_o + \Delta n\cos(kx)$, where n_o is the average density and Δn is the amplitude of the density fluctuation. Finally, we treat the ions as fixed on the high-frequency time scale and describe the electrons as fluid with density n_e, mean velocity u_e, and pressure p_e.

To derive an equation for the high-frequency electron density fluctuations, we start with taking a time derivative of the 1-D continuity equation,

$$\frac{\partial^2 n_e}{\partial t^2} + \frac{\partial^2}{\partial x \partial t}(n_e u_e) = 0 \tag{111}$$

Expand the last term and substitute $\partial n_e/\partial t$ by the continuity equation again. We will have,

$$\frac{\partial}{\partial x}\left(n_e\frac{\partial u_e}{\partial t}\right) = \frac{\partial}{\partial x}\left(u_e\frac{\partial(n_e u_e)}{\partial x}\right) - \frac{\partial^2 n_e}{\partial t^2} \tag{112}$$

Recall the 1-D momentum equation for a fluid,

$$m_e n_e\frac{du_e}{dt} = -q_e n_e E - \frac{\partial P_e}{\partial x} \tag{113}$$

where P_e is the electron pressure and is proportional to $k_B T_e n_e$. Here we ignore collisional effect. Take a spatial derivative of the left-hand side of (113), we have

$$\begin{aligned}\frac{\partial}{\partial x}\left(n_e\frac{du_e}{dt}\right) &= \frac{\partial}{\partial x}\left[n_e\left(u_e\frac{\partial u_e}{\partial x}+\frac{\partial u_e}{\partial t}\right)\right] \\ &= \frac{\partial}{\partial x}\left(n_e u_e\frac{\partial u_e}{\partial x}\right)+\frac{\partial}{\partial x}\left[u_e\frac{\partial(n_e u_e)}{\partial x}\right]-\frac{\partial^2 n_e}{\partial t^2} \\ &= \frac{\partial^2}{\partial x^2}(n_e u_e^2)-\frac{\partial^2 n_e}{\partial t^2}\end{aligned} \tag{114}$$

Then combine with the spatial derivative of the right-hand side of (112) and rearrange, we get,

$$\frac{\partial^2 n_e}{\partial t^2}-\frac{\partial^2}{\partial x^2}(n_e u_e^2)-\frac{q_e}{m_e}\frac{\partial(n_e E)}{\partial x}-\frac{1}{m_e}\frac{\partial^2 P_e}{\partial x^2}=0 \tag{115}$$

Next, we linearize this equation by assuming $n_e = n_o + \Delta n\cos(kx) + \tilde{n}$, $E = E_1 + \tilde{E}$ and $u_e = u_o + \tilde{v}$, where $\tilde{n}$, $\tilde{E}$ and $\tilde{v}$ are small perturbations of the electrons and u_o is the drift velocity of electrons in the electric field E_1. We treat $\tilde{n} \ll \Delta n \ll n_o$ and use an adiabatic equation of state for the pressure, i.e. $P_e = 3k_B T_e n_e$, assuming $\omega/k \gg v_{te}$, where v_{te} is the electron thermal velocity. Then we substitute these three perturbed quantities into (115). The first term becomes $\partial^2\tilde{n}/\partial t^2$ since n_o and Δn are assumed to be independent of time. The second term is zero since the drift velocity u_o is set to zero. The third term is the most important. It shows coupling between the pump field E_1 and the ion density fluctuation. Here, we ignore the very small perturbation in the electron density so that we get $n_e E = n_o E_1 + \Delta n\cos(kx)E_1 + \tilde{E}n_o$. For the last term, since kv_{te} is much smaller than the wave frequency ω, where $v_{te}^2 = k_B T_e/m_e$, we only retain $\tilde{n}$.

$$\frac{\partial^2\tilde{n}}{\partial t^2}-\frac{q_e}{m_e}n_o\frac{\partial\tilde{E}}{\partial x}-3v_{te}^2\frac{\partial^2\tilde{n}}{\partial x^2}=-\frac{q_e E_1}{m_e}\Delta n\cdot k\sin(kx) \tag{116}$$

Then from Poisson's Equation, $\partial\tilde{E}/\partial x = q_e n_e/\epsilon_o$. Therefore we have an inhomogeneous partial differential equation in $\tilde{E}$.

$$\frac{\partial}{\partial x}\left[\frac{\partial^2\tilde{E}}{\partial t^2}+\omega_{pe}^2\tilde{E}-3v_{te}^2\frac{\partial^2\tilde{E}}{\partial x^2}\right]=\frac{q_e^2 E_1}{m_e\epsilon_o}\Delta n\cdot k\sin(kx) \tag{117}$$

Integrate the above equation with respect to x,

$$\left(\frac{\partial^2}{\partial t^2}+\omega_{pe}^2-3v_{te}^2\frac{\partial^2}{\partial x^2}\right)\tilde{E}=-\omega_{pe}^2\frac{\Delta n}{n_o}E_1\cos(kx) \tag{118}$$

This equation describes the excitation of an electron plasma wave by the interaction of the pump field with an ion density fluctuation.

Parametric instability

We can now demonstrate a simple qualitative example of parametric instability. An ion density fluctuation couples an electromagnetic wave into an electron plasma wave to give us $\tilde{E}$. In turn, the electron plasma wave beats with the electromagnetic wave to generate a spatial variation in the electric field intensity (that is $E_o(\mathbf{x})$ discussed in the first part of this section), which can enhance the ion density fluctuation via the pondermotive force. Hence, a feedback loop is formed and depending on the pump amplitude, instability can result. Such an instability is called parametric instability, the parameter being the amplitude of the wave. Detailed instability analysis of this case can be found in [10] Kruer.

In order for the parametric instability to occur, it requires a minimal set of common characteristics:

- Matching condition — The spatially varying electric field E_o which is resulted from beating of two waves requires a wave-number matching to produce sustaining instability. Mathematically,

$$\mathbf{k}_o = \mathbf{k}_i + \mathbf{k}_s \tag{119}$$

  where the subscripts "o","i" and "s" stand for pump, idler and signal, respectively. In our example, the electromagnetic wave is the pump, the electron plasma wave is the idler, and E_o is the signal. The wave-number matching condition is equivalent to the conservation of momentum. However, the parametric instability has to satisfy the conservation of energy to take place. This will translate into the frequency matching condition.

$$\omega_o = \omega_i + \omega_s \tag{120}$$

  In our example, we would expect a high-frequency electromagnetic wave (the pump) beats with a high-frequency electron plasma wave which has a slightly smaller frequency than the pump to generate the signal, sometimes called daughter wave, which has the difference frequency of the pump and the idler.

- Threshold — The instability can only occur when the pump amplitude exceeds a critical value in order to maintain the feedback growth.

Parametric instabilities can be found in many physical systems, such as child swing. The pump in fact produces modulation of some physical parameter in the system which has a natural oscillation frequency. In our example, the natural frequency of the system is the electron plasma frequency and the signal frequency is the ion acoustic frequency. So the pump has to set at the sum of these two frequencies to give rise to parametric instability. In general, the instability analysis

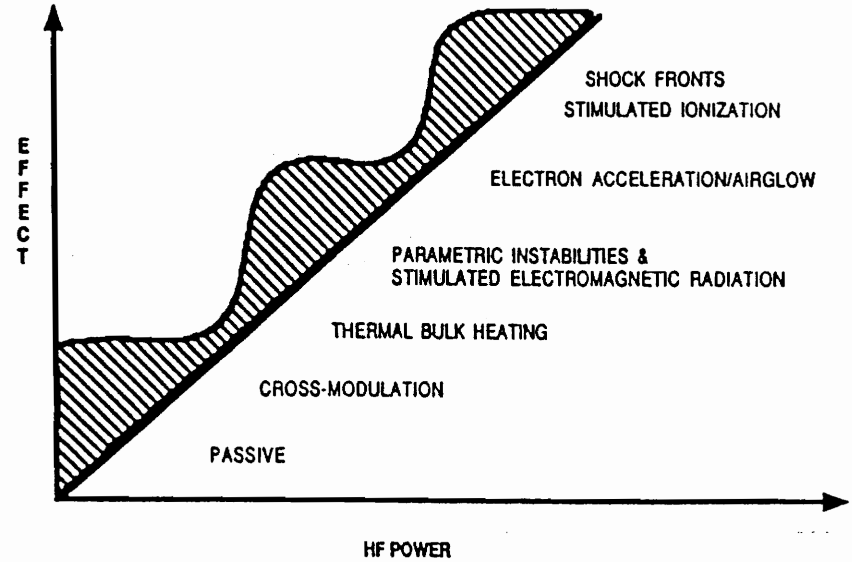


Figure 13: Hierarchy of heater thresholds in the ionospheric modification experiments (From [4] Carlson 1990).

requires techniqes in solving weakly nonlinear differential equations or weak tubulence analysis such as Mathieu equation ([17] Swanson, [12] Nicholson, and [3] Chen).

The example of ion density fluctuation reminds us about the similar mechanism employed in direct conversion which we discussed in the last section, although our example occurs in unmagnetized plasma. In both cases, an incident electromagnetic wave is converted into an electrostatic plasma wave through the matching conditions. However, in the direct conversion, the ion density fluctuation is provided by the atmospheric effects. That means it need not to have a critical pump amplitude to invoke the instability.

3.8 A survey of the proposed SEE theories

Since the discovery of stimulated electromagnetic emissions (SEE) in 1981, several published papers attempted to propose theories for some of the observed SEE features ([20] Stubbe 1984, [9] Leyser 1988, [12] Leyser 1991, [6] Goodman 1991, [18] Rao 1990, [19] Rao 1992 and [7] Huang 1993). Here we briefly summarize the essence of these proposed SEE theories. The electromagnetic emissions may, in principle, be generated by many plasma processes and originate in different height regions. The detected emissions are an integration over the entire ionospheric interaction region.

For the pump power used in Tromsϕ and Arecibo, parametric instabilities are likely candidates for the production of many SEE spectral features. Figure 13 shows the hierarchy of heater threshold power in relation to nonlinear processes that might occur during the ionospheric modification experiments. In fact, from the literature ([25] Wong 1987), we can have an estimate of electric field strength in experiments at Arecibo. The pump power is 400KW and the antenna gain is about 23dB. This corresponds to a wave electric field strength of up to 0.35V/m in the ionosphere at 200Km, without taking into account the standing wave pattern due to reflection. The up going and down going waves are in phase with each other at some locations near the reflection layer in the ionosphere, so that the resulting standing wave has maximum altitude. As discussed in section 3.4, the behavior near a plasma cutoff point can be modelled by an Airy equation. The solution is a sum of two independent Airy functions. The maxima of the solution are termed as Airy maxima. A plot of the wave electric field distribution can be found in [11] Leyser. The maximum field strength at the first Airy maximum (about 100m below the reflection point) is estimated to be ~3V/m at the incident power of 400KW ([25] Wong 1987 and [11] Leyser). This value well exceeds the threshold field strength of ~0.7V/m for the parametric decay instability ([24] Wong 1971). Enhancement of the wave electric field strength at Airy maxima due to the standing wave pattern is sometimes referred as swelling effect of the field. The standing wave pattern can also be seen in the 1-D electromagnetic simulation of plasma ([5] Close 1990).

Besides the swelling effect, enhancement of electrostatic plasma wave can happen at the upper hybrid layer through one of the mode conversion mechanisms discussed in section 3.6. This is also shown in the 1-D electromagnetic simulation ([5] Close 1990). The wave amplitude can be greatly enhanced so that the parametric instability threshold is exceeded and even a large cavity is formed ([25] Wong 1987). In this case, the mode converted electrostatic field is essentially perpendicular to the geomagnetic field ([11] Leyser).

The two above-mentioned enhancements of electric field reveal that there are possibly two major interaction regions responsible for different SEE features through parametric decay instabilities. Note that the nature of the two waves is different, one being electromagnetic and other being electrostatic. The continuum and DP are believed to be generated just below the reflection point, while the DM and BUM are produced at the upper hybrid layer. Parametric decay instability (PDI) is a favorate theory because of the asymmetry of almost all SEE features. The "decay" means the mother wave generates two daughter waves at lower frequencies. Sometimes, it is called a three-wave interaction or weak plasma tubulence theory. A summary of the theories proposed to explain the dominant SEE features are as follows.

1. Continuum and downshifted peak (DP)

Both Stubbe and Leyser proposed that the DP is produced by a single PDI and the continuum is a result of cascaded PDI of the same type ([20] Stubbe 1984 and [11] Leyser). The initial decay channel is that an electromagnetic wave at the first few Airy maxima decays into a Langmuir wave[7] and an ion acoustic wave[8].

$$\text{Standing EM wave} \underset{\text{PDI}}{\longrightarrow} \text{Langmuir wave} + \text{Ion acoustic wave}$$

The dispersion of Langmuir wave has been mentioned in section 3.5 and is repeated here.

$$\omega_e^2 = \omega_p^2 + 3k_e^2 v_{te}^2 = \omega_p^2(1 + 3k_e^2 \lambda_D^2) \tag{121}$$

where the subscript "e" refers to the electron plasma wave or Langmuir wave. The dispersion relation of an ion acoustic wave is given by ([12] Nicholson),

$$\omega_i^2 = k_i^2 c_s^2 \tag{122}$$

where the subscript "i" means ion-acoustic wave and c_s is the sound speed of a plasma, which is,

$$c_s = \sqrt{\frac{\gamma_e k_B T_e + \gamma_i k_B T_i}{m_i}} \tag{123}$$

Then the PDI occurs when the following matching conditions are satisfied.

$$\omega_o = \omega_e + \omega_i \tag{124}$$

$$\mathbf{k}_o = \mathbf{k}_e + \mathbf{k}_i \tag{125}$$

where the subscript "o" refers to the standing electromagnetic wave. Since the wavelength of the electromagnetic wave is much longer than the wavelengths of the parametrically excited electrostatic Langmuir and ion acoustic waves, the wave vector matching condition can be approximated as $\mathbf{k}_e + \mathbf{k}_i \approx 0$, i.e. the two daughter waves travel in opposite directions. Using the approximate matching conditions and the two dispersion relations, the Langmuir frequency can be solved. However, the above argument assumed an unperturbed density profile. In reality, the pondermotive force of the standing electromagnetic wave and the excited Langmuir and ion acoustic waves can significantly modify the density concentration in the interaction region. The electron depletion by pondermotive force can change the local plasma frequency and hence the Langmuir frequency. Thus, Leyser extended the above idea to include the pondermotive effect on the Langmuir frequency ([11] Leyser).

[7] It has been discussed in section 3.1 and 3.5. The dispersion relation of Langmuir wave is (91).

[8] It is a propagating electrostatic wave mode in a plasma. The propagation mechanism is analogue to a sound acoustic wave in a fluid. Unlike the sound wave, the longitudinal compression and rarefraction of ions are coupled through the Coulomb force. The dispersion relation is shown in (122). Details of ion acoustic wave can be found in most of plasma texts such as [3] Chen, [12] Nicholson and [17] Swanson.

The resulting Langmuir frequency is found to be around 1.5-2KHz below the pump frequency at the first four Airy maxima which are located at a few hundred meters below the reflection point. This frequency is about the same as the DP frequency. The Langmuir wave may be mode-converted back to an radiating electromagnetic wave at the same frequency because the density gradient of the ionosphere can facilitate this back conversion ([11] Leyser). Hence the initial PDI can account for the existence of the DP feature.

To account for the continuum, the successive decay of the above process is proposed ([20] Stubbe 1984 and [11] Leyser). The parametrically enhanced Langmuir wave can itself exceed the threshold for further parametric decay and excite another Langmuir and ion acoustic waves.

$$\text{Initial Langmuir wave} \underset{\text{PDI}}{\longrightarrow} \text{Second Langmuir wave} + \text{Ion acoustic wave}$$

If the process goes on, the injected pump wave can generate a wide spectrum of plasma waves. The enhanced Langmuir waves may contribute to the SEE spectrum, namely the continuum, provided that they can excite electromagnetic waves through linear conversion. The matching conditions for the successive decay is similar to the initial decay, except that the mother wave now is the initial Langmuir wave and hence the wave vector condition cannot be approximated as before.

$$\omega_e = \omega_{e'} + \omega_{i'} \tag{126}$$

$$\mathbf{k}_e = \mathbf{k}_{e'} + \mathbf{k}_{i'} \tag{127}$$

where the primes denote successive decay. Further decays are possible as long as the threshold for the PDI is exceeded. The frequency of the successively enhanced Langmuir waves are approximately given by ([11] Leyser),

$$\omega_{e'} \approx \omega_o - (2n-1)\omega_i \qquad \text{for} \;\; n = 1, 2, 3, \cdots \tag{128}$$

In summary, the continuum and DP are generated through the same PDI mechanism, but the DP is due to single decay and the continuum is due to successive decays. Their interaction region is several hundred meters below the reflection point.

2. Downshifted maximum (DM)

Stubbe, Leyser, and Huang *et al* proposed three different theories concerning the generation of the DM ([20] Stubbe 1984 and [11] Leyser). Again, the mechanism is PDI, but in the first and the latter two theories, the decay channels are different. Stubbe suggested the same PDI used in the generation of the continuum and DP, that is the decay of a standing electromagnetic wave into an ion acoustic wave and Langmuir wave cascaded by another similar decay, with a consideration of so called height spread effect[9] ([20] Stubbe 1984). The result is that the second Langmuir wave escapes from the ionosphere through the back conversion and contributes to the DM feature. However, a very weak point of Stubbe theory is that it cannot account for the sharp cutoff on the DM high-frequency side, which is approximately at the lower hybrid frequency (≈8KHz)[10] The involvement of lower hybrid frequency implies the importance of the geomagnetic field and hence the upper hybrid interaction region. This leads to Leyser's proposal of another theory occuring at the upper hybrid layer.

In Leyser's theory about the DM feature, it is assumed that an electrostatic wave (called the upper hybrid wave) at the upper hybrid layer is generated by the pump wave through one of the mode conversions discussed in section 3.6.

$$\text{O-mode EM wave} \xrightarrow[\text{Mode conversion}]{\hspace{4cm}} \text{Upper hybrid wave}$$

The electric field of the upper hybrid wave is essentially perpendicular to the magnetic field. An electromagnetic wave and a lower hybrid wave[11] are parametrically excited by the upper hybrid wave.

$$\text{Upper hybrid wave} \underset{\text{PDI}}{\longrightarrow} \text{O-mode EM wave} + \text{Lower hybrid wave}$$

The parametrically excited electromagnetic wave then contributes to the DM feature when it is received on the ground. The matching conditions are,

$$\omega_u = \omega_{DM} + \omega_l \tag{129}$$

$$\mathbf{k}_u = \mathbf{k}_{DM} + \mathbf{k}_l \tag{130}$$

where ω_u and ω_l are the frequencies of the upper hybrid and lower hybrid waves. They differ from the upper hybrid and lower hybrid frequencies discussed in section 3.4. Then Leyser used the two-fluid and kinetic models to solve for the growth rate of lower hybrid wave [12] ([11] Leyser). The mathematics and approximation are quite involved, so they are not repeated here. Leyser claims that since the growth rate of the lower hybrid wave is the highest at the frequencies slightly (~1KHz)

[9] It is due to the weight of the spectral components of the stimulated radiation with the altitude over which they are generated.

[10] At 200Km, the dominant ion species is O^+. The lower hybrid frequency f_{LH} is approximately given by (61) and is equal to $1.36MHz/\sqrt{1836 \times 16} = 7.93KHz$, where the electron cyclotron frequency is taken to be 1.36MHz.

[11] It is an electrostatic ion wave propagating perpendicular to the magnetic field. The dispersion relation of a lower hybrid wave is $\omega^2 = k^2c_s^2 + |\Omega_i\Omega_e| = k^2c_s^2 + \Omega_{LH}^2$, where c_s is same as in (123). More details can be found in [12] Nicholson and [17] Swanson.

[12] Note that Leyser used a different dispersion relation for the lower hybrid wave in the two-fluid theory from the one quoted from [12] Nicholson.

above the lower hybrid frequency from the kinetic model, and it drops rapidly around 11KHz, it can account for the spectral shape of the DM feature.

Huang *et al* suggest that at upper hybrid layer, an electrostatic wave is produced by another mechanism other than the mode conversions discussed in section 3.6 ([7] Huang 1993). The thermal oscillating two stream instability (OSTI) leads to the parametric excitation of electron Bernstein and/or upper hybrid waves together with purely growing density irregularities by the O-mode pump wave. This is a four-wave interaction process. The matching conditions for the thermal OTSI are,

$$\mathbf{k}_1 + \mathbf{k}_s \quad = 0 = \quad \mathbf{k}_2 - \mathbf{k}_s \tag{131}$$

$$\omega_1 + \omega_s^* \quad = \omega_o = \quad \omega_2 - \omega_s \tag{132}$$

where the subscripts "1" and "2" stand for electron Bernstein waves and/or upper hybrid waves, and "s" for the purely growing density irregularities. Note that the frequency $\omega_s = i\gamma$ is purely imaginary and γ is the growth rate of the density irregularities. In their paper, only the third electron cyclotron harmonic is considered. Their analysis shows that the growth rate of the upper hybrid wave is strong slightly below the upper hybrid layer when the pump frequency is less than the third harmonic of electron cyclotron frequency. The generation mechanism of upper hybrid waves is,

$$\text{O-mode EM wave} \xrightarrow[\text{Thermal OTSI}]{\hspace{3cm}} \text{Upper hybrid wave}$$

Therefore the excited upper hybrid waves are essentially localized in the region slight below the upper hybrid layer. The downshited maximum feature is produced by the same parametric decay instability proposed by Leyser.

3. Upshifted maximum (UM)

Stubbe suggested two parametric instability mechanisms to account for the UM feature ([20] Stubbe 1984). In this case, the parametric decay instability cannot by itself facilitate the generation of UM that the frequency of a daughter wave is higher than the pump. Instead, at least one parametric instability (PI) has to shift up the frequency of one of the daughter wave. Stubbe proposed two possible routes as follows.

(a) A Langmuir wave is initially generated through the parametric instability as follows.

$$\text{Standing EM wave} \xrightarrow[\text{PI}]{} \text{Langmuir wave} + \text{Ion acoustic wave}$$

The Langmuir frequency is shifted up to $\omega_e = \omega_o + \omega_i$ (n=1). The parametrically excited Langmuir wave then decays into the second Langmuir wave and an ion acoustic wave through a cascaded parametric instability.

$$\text{Initial Langmuir wave} \underset{\text{PI}}{\longrightarrow} \text{Second Langmuir wave} + \text{Ion acoustic wave}$$

The frequency of the excited Langmuir frequency is $\omega_e = \omega_o + 3\omega_i$ (n=2) The final stage is to shift down the above frequency by the PDI.

$$\text{Second Langmuir wave} \underset{\text{PDI}}{\longrightarrow} \text{Third Langmuir wave} + \text{Ion acoustic wave}$$

The frequency of the last Langmuir wave is $\omega_e = \omega_o + 2\omega_i$, which corresponds to the frequency of UM, provided that it is converted to an electromagnetic wave through mode conversion.

(b) The second possibility involves a two-stage parametric instability. The standing electromagnetic wave decays into an ion acoustic and Langmuir waves as follows.

$$\text{Standing EM wave} \underset{\text{PI}}{\longrightarrow} \text{Langmuir wave} + \text{Ion acoustic wave}$$

This is the same process as in the first case. The Langmuir frequency is shifted to $\omega_e = \omega_o + \omega_i$. Then the Langmuir wave decays into an electromagnetic wave and an ion acoustic wave.

$$\text{Initial Langmuir wave} \underset{\text{PI}}{\longrightarrow} \text{EM wave} + \text{Ion acoustic wave}$$

The frequency of the radiating electromagnetic field is shifted up again by an ion acoustic frequency to $\omega_e = \omega_o + 2\omega_i$.

Stubbe also argues that the UM feature is weak because these decay processes require the plasma to supply extra energy to shift up the stimulated frequency.

4. Broad upshifted maximum (BUM)

There exist two theories about the BUM feature proposed by Leyser and Goodman. From the empirical frequency relation of the BUM as in section 2.4/footnoteThat is $f_{BUM} = 2f_o - nf_{ce}$., Leyser suggests that the BUM may be caused by a four-wave interaction [13] ([11] Leyser). The four waves are two pump photons or upper hybrid plasmons, a decay mode at nf_{ce}, and the stimulated radiation at f_{BUM}.

On the other hand, Goodman attempted to calculate a current that generates the BUM feature ([6] Goodman). The incident electromagnetic field with a wave frequency slightly above the electron

[13] See [12] Nicholson.

cyclotron harmonic excites a current in the ionospheric plasma. Then this current continues to excite electron Bernstein waves[14] and a lower hybrid wave. The resulting current, which depends on density gradients across the geomagnetic field, radiates an electromagetic field at the BUM frequencies to the ground. It can reproduce the empirical frequency relation of the BUM and there is a cutoff in the spectrum at about 10KHz above the pump frequency, which is observed from the experimental SEE spectra.

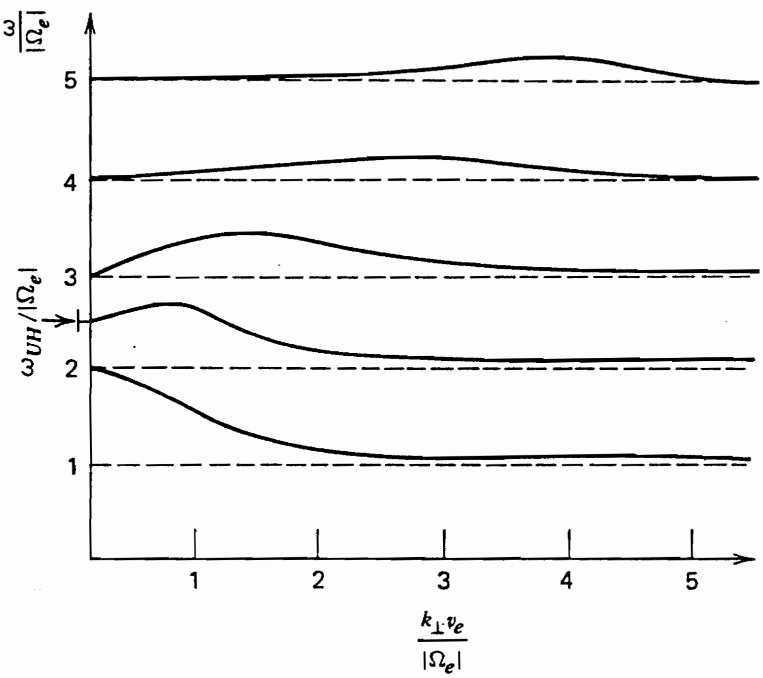


Figure 14: Sketch of the dispersion curves for electron Bernstein modes (From [12] Nicholson).

5. Broad symmetrical structure (BSS)

So far there is no extended theoretical proposal other than the report paper on the BSS ([21] Stubbe 1990). They suggest that the electron Bernstein modes should come into play because the

[14]Bernstein modes are new wave modes which can only be derived using kinetic formulation. These waves depend on the detailed interaction of the wave motion with the gyro-orbits of the particles. The derivation is tedious and can be found in many plasma texts such as [16] Stix, [17] Swanson, [12] Nicholson, and [3] Chen. The dispersion curves of electron Bernstein modes are shown in Figure 14. These modes propagate across the magnetic field. Note that for frequencies above the upper hybrid frequency, there are stop bands where no wave can exist. There also has ion Bernstein modes. Its dispersion curves are exactly the same as the electron Bernstein modes, except that the electron cyclotron frequency Ω_e and the upper-hybrid frequency ω_{UH} are replaced by the ion cyclotron frequency Ω_i and the lower-hybrid frequency ω_{LH}, respectively.

BSS occurs only in a narrow frequency range around the third electron cyclotron harmonic. Again, parametric instabilities are proposed to account for the generation of the BSS. Lower hybrid waves may be generated by parametric decay of primary Bernstein-upper hybrid waves of frequency f_o into secondary Bernstein-upper hybrid waves and lower hybrid waves. The Bernstein-upper hybrid waves are electron Bernstein waves with the upper hybrid frequency very close to the electron cyclotron frequency. The BSS may be understood as being due to scattering of primary Bernstein-upper hybrid waves by lower hybrid waves having the same or the opposite propagation direction. The secondary electromagnetic waves generated in this way would possess a spectrum which is symmetric around f_o.

6. Quenching of DM

There are three competing theories proposed by Leyser, Rao and Kaup, and Huang *et al.* Leyser suggests that electron cyclotron damping [15] quenches the mode converted upper hybrid waves to a very low amplitude so that the parametric decay instability cannot be triggered, when the pump frequency is close to the electron cyclotron harmonics. The complete quenching of DM has been observed in one of the recent SEE experiment only within a 200Hz bandwidth around the seventh harmonic ([13] Leyser 1992). Hence Leyser claims that the complete quenching of DM is evident from electron cyclotron damping and can be used to measure the local magnetic field strength with high accuracy.

Rao and Kaup suggest that the mode conversion of an upper hybrid wave into electron Bernstein waves can be the mechanism responsible for the sudden quenching of the DM feature in the SEE experiments ([18] Rao 1990). They also claim that for upper hybrid waves propagating exactly orthogonal to the ambient magnetic field, there is no cyclotron damping since the particle motion is elliptic in the perpendicular plane and hence the particles cannot keep in phase with the wave. Damping arises if the wave has a small wave number ($k_{\|}$) parallel to the magnetic field direction such that the resonance condition $\omega - n\Omega_e = k_{\|}v_{\|}$ is satisfied. For the Tromsϕ experiment, it can be shown that the resonance condition is quite stringent to be satisfied unless the pump frequency is extremely close to $n\Omega_e$. Hence they propose that when $f_o \approx nf_{ce}$, for $n \geq 3$, the upper hybrid waves can be efficiently mode-converted into nonpropagating electrostatic electron Bernstein modes. In a later paper ([19] Rao 1992), they calculate the bandwidth around the third, fourth and fifth

[15] Cyclotron damping is a kinetic effect which is similar to Bernstein modes. The difference is that cyclotron damping occurs when the particle sees a wave whose Doppler frequency is its cyclotron harmonics : $\omega - k_{\|}v_{\|} = n\Omega_j$ for $n = \pm1, \pm2, \cdots$. The particle is then continuously accelerated and the wave is damped. An example of cyclotron damping is shown in Figure 15, where the pump frequency is at twice of the cyclotron frequency. The particle is accelerated twice in one revolution when the peaks of the electric field aligns the particle velocity at $\Omega t = 0$ and $\Omega t = \pi$.

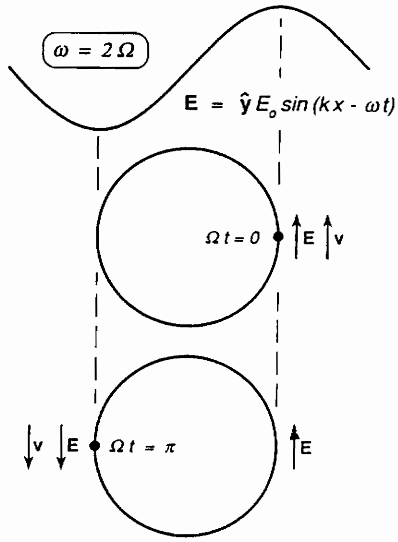


Figure 15: Illustration of cyclotron damping at $\omega = 2\Omega$ (From [16] Stix).

harmonics of electron cyclotron frequency that the upper hybrid wave undergoes mode conversion into electrostatic Bernstein modes. The bandwidth is maximum ($\approx 14KHz$) for the third harmonic, but decreases very rapidly for for the higher harmonics.

Huang *et al* propose that the quenching of DM near $3f_{ce}$ is because besides the upper hybrid wave, the nonpropagating electron Bernstein waves are excited through the thermal OTSI ([7] Huang 1993). When the heater frequency is slightly higher than $3f_{ce}$, the upper hybrid wave and the electron Bernstein wave become linearly coupled and hence the growth rate of the upper hybrid wave is much smaller than the case when the pump frequency is less than $3f_{ce}$. Furthermore, the height region of exciting the thermal OTSI below the upper hybrid resonance layer shrinks as the heater frequency approaches $3f_{ce}$. The net result is that the amplitude of the upper hybrid wave decreases rapidly when the pump frequency is very close to $3f_{ce}$ so that the PDI cannot be triggered to generate the DM feature.

All of the proposed SEE theories we have described are in general heuristic. At this time, there is no unified and definitive theory that explains the SEE spectrum in detail. Some of these theories have obvious weaknesses and inconsistencies with the experimental observations. We will now discuss some of these difficulties.

- Downshifted peak (DP) — In Tromsϕ experiments, the standing wave electric field is nearly perpendicular to the geomagnetic field. Although Leyser claims that the electric field is parallel to the magnetic field ([11] Leyser), it is difficult to conceive how a nearly perpendicular electric field can be transformed into a parallel one without losing its electromagnetic character. His theory will be more reasonable if the standing electromagnetic wave is somewhat mode converted into an electrostatic wave parallel to the geomagnetic field, which in turn undergoes the same PDI to generate the DP. Otherwise, the ambient magnetic field has to be taken into account. Moreover, Leyser's model cannot explain why the DP feature happens intermittently around the third harmonic of electron cyclotron frequency and why it is favored when the pump frequency is far below the critical frequency.

- The continuum — Since the continuum is generated by the same PDI mechanism of the DP followed by successive decay processes, besides the above-mentioned difficulties concerning the DP, one may suspect why the DP feature does not always show up with the continuum because the initial PDI has to occur before the subsequent decay processes can start. This is best demonstrated by the fact that at a very low pump power, the continuum is the only feature and the DP never appears. Moreover, the frequency coverage of the continuum typically extends to 15KHz. Suppose the ion acoustic frequency is about 2KHz. The number of successive decays is about $15/(2 \times 2) \approx 4$ and the frequency step between cascaded decays is 4KHz. Thus, one should expect to see at least four (including the DP if it exists) discrete peaks on the continuum envelop. In fact, this is an ideal case because the last few decays may not be triggered because the threshold cannot be exceeded. In this case, one may expect to see an abrupt cutoff on the left edge of the continuum, which is never observed. All these arguments suggest that the continuum may be generated by other physical processes which are different from that generates the DP.

- Downshifted maximum (DM) — Leyser's theory is generally more acceptable than Stubbe's because it can explain the cutoff on the high-frequency side of the DM. But, it does not consider how the parametrically excited electromagnetic wave can propagate to the ground receiver. This electromagnetic wave is parametrically generated in the upper hybrid region where the pump frequency is approximately equal to the upper hybrid frequency. The frequency of the DM, f_{DM}, is typically 8 to 10KHz lower than the pump. Thus, it has to go through a region where the upper hybrid frequency equals f_{DM}. Strong absorption and/or reflection may occur. Also, similar to the DP feature, Leyser's proposal cannot provide answers for the following facts. The DM is favored when the pump frequency is close to nf_{ce} or when the pump frequency is near the critical frequency of the ionosphere[16], and the DM sometimes disappears from the

[16] It is quite ambiguous to justify whether the condition that the pump frequency is near the critical frequency favors

SEE spectrum. Also, the theory cannot predict the frequency dependence of the DM on the pump frequency, according to the empirical formula discussed in section 2.4.

- Upshifted maximum (UM) — Stubbe's theory is not viable because it cannot account for the fact that the offset frequency of the UM from the pump is always less than that of the DM. In fact, as will be seen from the simulation result, the UM frequency is related to the lower hybrid wave, that is dependent on the magnetic field. It is more likely generated from the upper hybrid region rather than the reflection region which is proposed by Stubbe.

- Broad upshifted maximum (BUM) — Goodman's theory is difficult to verify whether it is reasonable or not because only small amount of numerical evaluation of the theory can be found in the original paper ([6] Goodman). On the other hand, Leyser's conjecture of four-wave interaction seems viable. But the frequency of one of the four waves is at nf_{ce}. A frequency component at this harmonic frequency should show up in the SEE spectrum together with the BUM. Of course, it is not the case in the experimental spectrum. The second point is that a plasma wave at nf_{ce} should undergo cyclotron damping unless it propagates exactly parallel to the magnetic field. Also, as have been pointed out in section 2.4, the frequency relation of the BUM, which leads Leyser to suggest the four-wave interaction, is a very rough estimate because the BUM shape is highly variable.

- Broad symmetrical structure (BSS) — It is not a mature time to comment anything on Stubbe and Kopka's proposal. However, it should be noted that the BSS is somewhat related to the DP because they possess similarity that they appear intermittently in a narrow range of frequencies around $3f_{ce}$. Furthermore, the BSS never coexists with the DP. It is logical to presume that the BSS is also generated from the reflection layer, rather than the upper hybrid layer suggested by Stubbe and Kopka. In fact, as will be discussed in the next chapter, the BSS feature has never been observed in the simulation of the upper hybrid layer.

- Quenching of DM around nf_{ce} — Amongst the three proposed theories, the author believes that the two theories proposed by Leyser, and Rao and Kaup seem to be more reasonable. Although Leyser's cyclotron damping theory is commented that the parallel wave vector $k_{\|}$ has to exist for strong cyclotron damping when the pump frequency is not exactly at nf_{ce}, it may still be one of the quenching mechanisms of DM because it is only necessary for the upper hybrid wave to be damping to a level where the PDI threshold is not exceeded. Hence cyclotron damping may have effects on the quenching of DM even though the pump frequency is slightly

the development of DM because it is difficult to tell whether it actually means that many DMs occur or a strong DM appears in this case.

offset from the exact electron cyclotron harmonics. In fact, one cannot conclude accurately about the quenching range of DM for the third, fourth and fifth harmonics from the displayed data in [11] Leyser because Leyser did not conduct a dedicated experiment similar to that performed in Russia for the seventh harmonic, to measure the quenching range. Unlike the seventh harmonic, it may encounter a difficulty that the frequent appearance of the continuum disturbs an accurate measurement of such a range since the spectra of the continuum and the DM often overlap. When compared with the Rao and Kaup theory, cyclotron damping differs from the mode conversion of upper hybrid wave into electron Bernstein modes in the fact that the quenching range of cyclotron damping includes the frequencies slightly below nf_{ce}, but electron Bernstein modes do not. Hence one can design a series of plasma simulations within a narrow range of frequencies around nf_{ce} to determine which one of the proposed mechanisms is crucially responsible for the quenching of DM. It is also possible that both take part in the quenching process.

Huang *et al* attempt to explain the quenching of DM by proposing another route that the upper hybrid wave is generated instead of mode conversions. Since the mode conversions are generally accepted and verified as a valid mechanism for the generation of an electrostatic wave at the upper hybrid layer, the proposed thermal OTSI has to show its significance over mode conversions before it can be claimed for the dominant mechanism responsible for the quenching of DM. Another difficulty of this newly proposed theory is that they claim the upper hybrid wave is generated in the region slightly below the upper hybrid layer. But they did not state the exact height where the upper hybrid wave is the strongest in their paper. It is natural to have an upper hybrid wave generated at an altitude where the pump frequency is equal to the upper hybrid frequency because the upper hybrid layer is a resonance for the pump wave. The wave vector and hence the wave velocity of the upper hybrid wave is small so that it can stay at the upper hybrid layer for a long time to generate the SEE. Hence, in the thermal OTSI theory, it is reasonable to expect that the parametrically excited upper hybrid wave has a smaller frequency than the pump in order for the excited wave to stay in the regions slightly below the upper hybrid layer. From the data shown in [7] Huang, if the electron cyclotron frequency is assumed to be 1.36MHz, at $f_o = 4.05MHz$, the height range which is in terms of the plasma frequency range, is calculated to be $\Delta f_{pe} = 8.16KHz$. Since $f_{UH}^2 = f_{pe}^2 + f_{ce}^2$, for small changes in both frequencies, $\Delta f_{UH} \approx \Delta f_{pe}(f_{pe}/f_{UH})$. Suppose the upper hybrid waves are needed to have frequencies equal to the local upper hybrid frequencies in order to stay in the regions below the upper hybrid layer. For 4.05MHz, the maximum deviation of the local upper hybrid frequency from the pump frequency is $8.16 \times 3.81/4.05 = 7.68KHz$. Since Huang *et al* suppose the same PDI to generate the DM, the frequencies of the DM would at least extend from the lower hybrid frequency (cutoff point) to the sum $(f_{LH} + 7.68KHz)$ below the pump frequency. As the pump frequency goes further below the third harmonic, according to the proposed thermal OTSI theory, the DM should be more spread out and stronger. This contradicts the experimental observation. If one compares the mode conversions and the thermal OTSI, the former channel is more direct and believable for the generation of the upper hybrid wave which in turn produces the DM feature.

Generally speaking, up to now, none of these theories are comprehensive and complete. In the following chapters, the simulation results will help us to decide whether some of the SEE theories are viable or not.

4 1-D numerical simulation

In the previous chapter, we studied the basic theory of waves in plasmas. Only in very simple cases can we completely work out the problem analytically. In most cases, we settle with linearizing the resulting partial differential equations in order to obtain an analytic solution for the propagation of small amplitude waves. Linearization essentially throws away all nonlinear properties inside the plasma. As an example of the inadequency of linear theory, consider the parametric instability analysis. We can predict whether the instability grows or not from linear theory. But, we cannot state how the instability grows and when it will saturate. When we deal with complicated processes such as SEE, a number of nonlinear physical processes occurs inside an inhomogeneous plasma. It is probably impossible to derive the end results that describe the system analytically without sacrifying the nonlinear details. Numerical simulation offers an alternative way to study the problem. From the simulation, we can pick out some of the dominant physical processes which are responsible for the generation of SEE. We can also investigate almost every detail (e.g. the phase space of a species at a particular time instant) in the plasma. Another advantage of computer simulation of plasma over experiments is that one can freely change all physical parameters. This sometimes facilitates the recognition of the most important underlying mechanisms. We devote this chapter to discuss the techniques of 1-dimensional (1-D) numerical simulation of plasmas.

4.1 Types of numerical simulation of plasmas

The types of numerical simulation of plasma have already been dictated in section 3.5 of Chapter 3 when we discussed the kinetic and fluid descriptions. In general, there are four major catagories of plasma simulation.

- Particle simulation.
- Vlasov or Boltzmann simulation.
- Fluid or MHD simulation.
- Hybrid simulation.

The most primitive way to simulate a plasma is to follow every particle in the plasma. This type of simulation is referred to as a particle code or particle simulation. Of course, one has to assign the initial position and velocity of each particle. The simulation program then uses the discretized versions of Newton's second law, the Lorentz force equation and Maxwell's Equations to calculate the subsequent motion of the particles. There are two subclasses of particle simulation: particle-particle

(PP) simulation and particle-in-cell (PIC) simulation. The PIC plasma simulation is sometimes called particle-mesh simulation. Their main difference is in the calculation of the force on each particle. This is best illustrated by an example. Let us consider an electrostatic plasma simulation. That means we ignore all magnetic fields generated by the motion of the charged particles. Sometimes this is called the electrostatic approximation to the full electromagnetic treatment. Electrostatic approximation is valid when the ratio of the particle pressure to the magnetic pressure, generally denoted by β[1], is much less than one. Such a plasma is called a low beta plasma. In an electrostatic simulation, we only need to consider Gauss's law in Maxwell's Equations because the other three are automatically satisfied (all zero) under this assumption. There are two ways to calculate the local electric field at a particle point. One way using Coulomb's law is to sum up all force contributions from each particle in the plasma. This is the PP method. Another way is to use Poisson's Equation to obtain the electric potential and then the electric field. Here we need to calculate the charge density. Since for a point charge, the charge density at the particle point is infinite, we have to evaluate the charge density macroscopically. That requires some form of meshes or grid cells to collect charge particles and then the charge density can be found by averaging all these grouped charges over a cell. After numerically solving Poisson's Equation, interpolation is required to distribute calculated force to every particle in every cell. This is the PIC method. For large number of particles, the PIC method is faster than the PP method in the calculation of force because the numerical solution of Poisson's Equation can utilize the fast Fourier Transform (FFT). We will discuss in more detail about the 1-D PIC electrostatic simulation later. Two classic books on particle simulation are [7] Hockey and [1] Birdsall.

The second type of numerical plasma simulation is the Vlasov code. It is essentially a numerical program solving a system of nonlinear partial differential equations, namely, Vlasov and Maxwell's Equations. This approach avoids statistical errors present in particle simulation, and has been used successfully. Note that there are seven independent variables in the Vlasov-Maxwell formulation. The numerical solution to these seven nonlinear differential-integral equations (Vlasov equation, four Maxwell's Equations, and two from charge and current densities) is not trivial and is prone to numerical instability due to the convective terms in the Vlasov equation. Therefore, numerical simulation of Vlasov equation is limited to a small number of dimensions and short time scale. It is possible to include collisional effects in the Vlasov equation. This is referred to as the collisional Vlasov equation or Boltzmann's equation. The collisional term is usually evaluated using Fokker-Planck equation. Thus this type of simulation is sometimes referred as Vlasov-Fokker-Planck code.

The next numerical plasma simulation we want to discuss is the fluid codes. Such a program

[1] The ratio β is defined as the ratio of the particle pressure $\sum nk_BT$ to the magnetic pressure $B^2/2\mu_o$ ([3] Chen).

simply solves the system of equations in the fluid formulation numerically. Here we have to be careful about the applicable regime of the fluid description of a plasma. Although the fluid formulation has only four independent variables, the convective terms are still present. Therefore, numerical stability becomes a consideration. Compared with Vlasov codes, this type of simulation is less complicated, though nontrivial, at the expense of losing information on kinetic behavior of the plasma. In some dense plasmas, the many-fluid model can be replaced by the magnetohydrodynamic formulation. One can then use MHD codes to simulate the plasma.

The last simulation type is a combination of fluid and particle simulation. Thus it is called a hybrid code. In some cases, the kinetic effect is crucial for only one of the species, for example, electrons in Landau damping. In such cases, one can use particle model for electrons and fluid model for ions because fluid codes is usually more time-efficient and less noisy than particle codes.

The particle-in-cell simulation is chosen to study the stimulated electromagnetic emissions. The reasons are two-fold. First, the time and space scales of interest is sufficiently large so that the Vlasov simulation can have severe stability problem. Moreover, the interactions between the excitation and particles and between particles are important in this problem. Only the particle code can essentially simulate these effects and reproduces the required information because the fluid code and the hybrid code do not simulate both interactions simultaneously. The PIC method is used because of its better computational efficiency discussed previously.

Plasma simulation in fact is a computer experiment to study plasma problems. It can discover new physical processes which may be difficult to do in laboratory experiments. In the future, theory, numerical simulation and experiments will be closely related together in order to develop new plasma physics.

4.2 1-D particle-in-cell electrostatic plasma simulation

In this section, we briefly discuss the one-dimensional particle-in-cell electrostatic plasma simulation program, generally known as ES1 ([1] Birdsall). One spatial dimension (x) is assumed and in the original version of ES1, two other velocities (v_x and v_y) are used. The whole plasma length (l) is equally divided into a number of grid cells (ng) which is required to be an integer power of two since FFT techniques are used. Thus, there are ng+1 grid points. We generally use the index i to denote particles and the index j to denote grid cells or grid points. Figure 16 gives a view how the 1-D geometry is divided. The parameter dx is obviously equal to l/ng.

The algorithm of ES1 is rather simple, as shown in Figure 16. The computation cycle consists of three major subroutines (MOVE, ACCEL and FIELD) to accomplish the jobs in each box. After each cycle, the program advances to the next time step. When the pre-defined total number of time

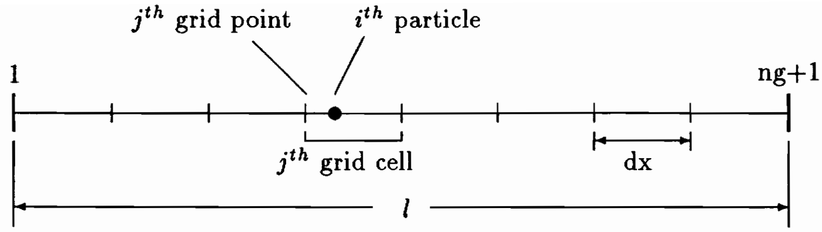


Figure 16: Discretization of the plasma length and naming of grids and particles in ES1.

steps (nt) is reached, the program ends and plots simulation histories for diagnostic purposes.

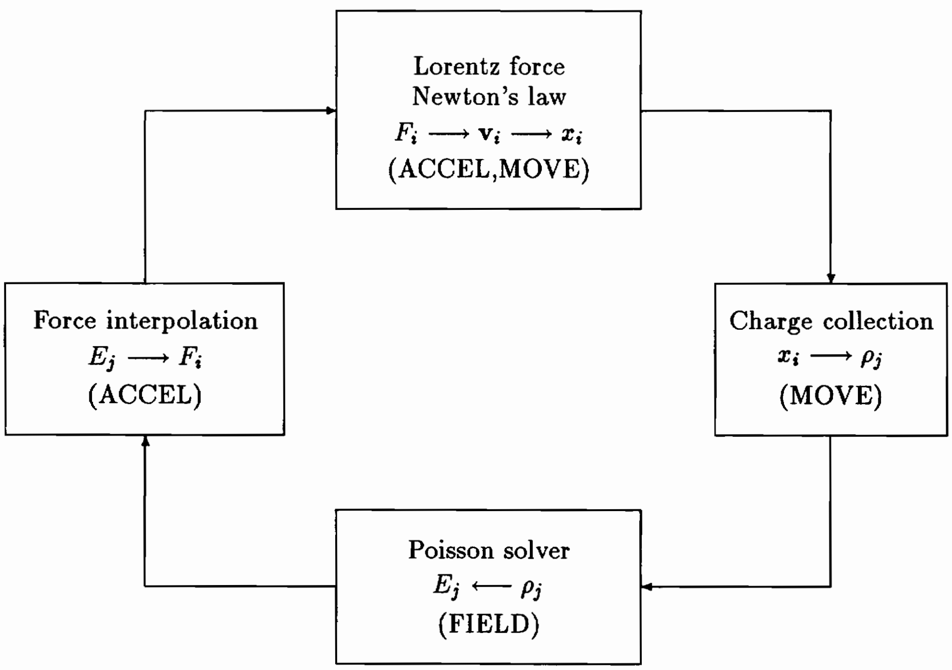


Figure 17: Basic algorithm of ES1.

Three equations, namely Newton's law, Lorentz force equation and Poisson's Equation, are used in the computation loop.

- Newton's second law : $\mathbf{F} = \mathrm{m}\frac{\mathrm{d}\mathbf{v}}{\mathrm{dt}}, \quad \mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{dt}}.$
- Lorentz force equation : $\mathbf{F} = \mathrm{q}(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$
- Poisson's Equation : $\nabla^2\phi = -\frac{\rho}{\epsilon_o}.$

Of course, we only keep the spatial dimension x in ES1. However, the velocities other than x-direction certainly have contribution to the force acting on particles through the Lorentz force equation. Therefore, we have to keep more than one velocity when the background magnetic field is nonzero. The first two equations are solved numerically using the center difference method, while the Poisson's equation is solved by fast Fourier Transform (FFT). All of them will be discussed in more detail in the following sections.

4.2.1 A physical picture

Figure 18 shows a physical picture of how a particle moves in 1-D electrostatic plasma simulation. The original particle is in 3-dimensional space. But in our 1-D program, we only keep track of the position x, that is an orthogonal projection of the particle position on the x-axis. When the magnetic field is perpendicular to the z-axis, the only velocities that can affect the particle motion in x-direction are v_x and v_y, through the Lorentz force equation. The original version of ES1 assumes this case. In this research, all the three velocities are implemented and the background magnetic field can be set at arbitrary angles. Note that the magnetic field is uniform over the whole plasma length.

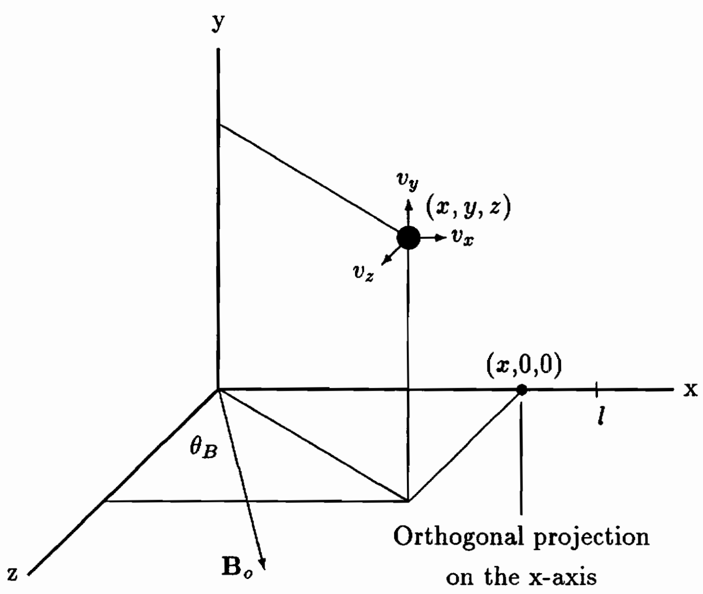


Figure 18: A physical picture of 1D 3V electrostatic plasma simulation.

4.2.2 Equation of motion (Lorentz force equation)

In Newton's second law, if we want to get a numerical solution to the position x, we have to integrate twice. Finite-difference methods are used in ES1 to calculate the velocities and position due to its simplicity and acceptable accuracy. The numerical procedure must be computationally efficient because a problem may call for more than 10000 particles to be run for more than 1000 time steps. Every computation cycle will go through the equation of motion.

The finite-difference method is based on the Taylor series expansion of a function about a particular point. That is,

$$u(x_o + h) = u(x_o) + \frac{h}{1!}\left(\frac{du}{dx}\right)_{x_o} + \frac{h^2}{2!}\left(\frac{d^2u}{dx^2}\right)_{x_o} + \cdots \tag{133}$$

where h is the finite step. The first derivative is then approximated by,

$$\left(\frac{du}{dx}\right)_{x_o} \approx \frac{u(x_o + h) - u(x_o)}{h} + O(h) \tag{134}$$

This is known as forward difference method. Similarly, one can change h to -h to obtain the backward difference equation.

$$u(x_o - h) = u(x_o) - \frac{h}{1!}\left(\frac{du}{dx}\right)_{x_o} + \frac{h^2}{2!}\left(\frac{d^2u}{dx^2}\right)_{x_o} + \cdots \tag{135}$$

$$\left(\frac{du}{dx}\right)_{x_o} \approx \frac{u(x_o) - u(x_o + h)}{h} + O(h) \tag{136}$$

However, if we subtract the Taylor representations of the forward and backward expansions about x_o, a more accurate formula for the first derivative can be obtained.

$$\left(\frac{du}{dx}\right)_{x_o} \approx \frac{u(x_o + h) - u(x_o - h)}{2h} + O(h^2) \tag{137}$$

This is called the center difference method. The higher accuracy comes from the fact that three points are simultaneously used. The center difference method is used in ES1 to solve for the velocities and position ([1] Birdsall). The update procedures of the velocities and position are $\frac{1}{2}\Delta t$ out of step[2]. This is often called the leap-frog scheme as shown in Figure 19. Hence the first set of equations of motion are,

$$m\frac{\mathbf{v}_{new} - \mathbf{v}_{old}}{\Delta t} = \mathbf{F}_{old} \tag{138}$$

$$\frac{x_{new} - x_{old}}{\Delta t} = v_{x(new)} \tag{139}$$

The leap-frog method has been found to be very accurate. When $\omega_{pe}\Delta t \leq 2$, there is no amplitude error in the simulation of high-frequency plasma oscillations ([1] Birdsall). But this does not imply there is no phase error. It may still be significant.

When the magnetic field is present in the simulation, the computation of the gyromotion of particles needs a particular numerical scheme to be compatible with the linear motion. This scheme is known as Boris mover ([1] Birdsall). The principle of Boris scheme is demonstrated using two velocities as follows. Consider the magnetic field along the z-direction ($\theta_B = 0$). Thus the gyromotion of all charged particles is in the x-y plane. The Lorentz force has two parts : one from the electric

[2] Here, $\frac{1}{2}\Delta t = h$, where Δt is the time step of the plasma simulation.

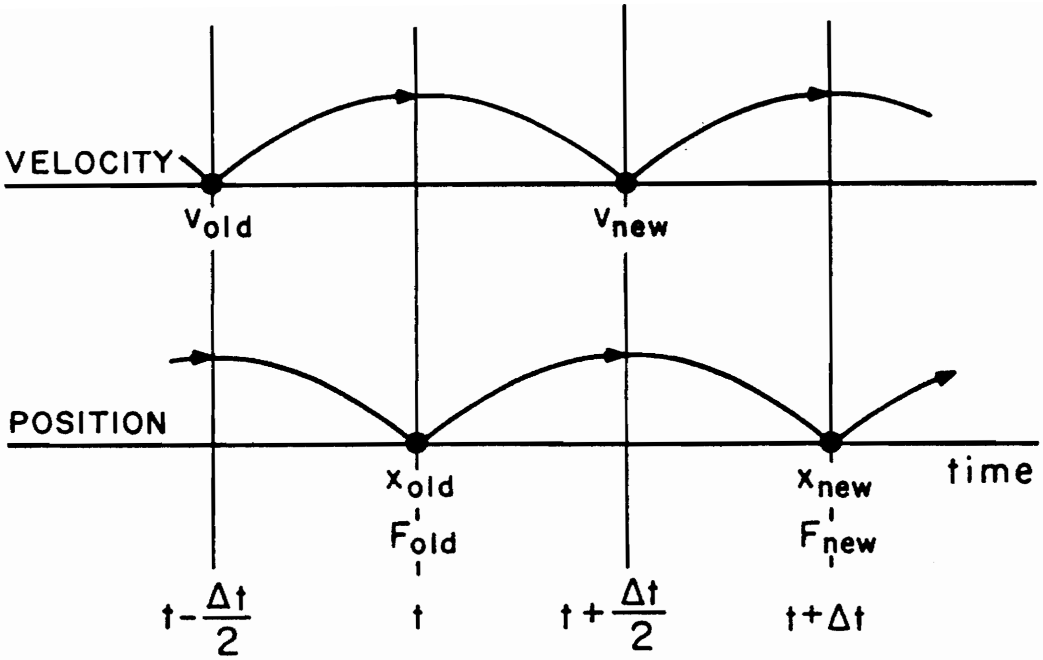


Figure 19: Leap-frog integration method used in ES1 (From [1] Birdsall).

field and the other from the magnetic field. Here, the electric field and magnetic field are to be calculated at the particle point. Using a spatial grid, we must interpolate the electric and magnetic fields from the grid to the particle point. The detailed weighting method will be discussed together with the Poisson solver in the next section.

From section 3.2, the kinetic energy of the gyromotion of a charged particle in a uniform magnetic field is constant. That means the magnitude of the velocity **v** responsible for the rotation is constant. The only possibility to change the magnitude of **v** is through v_x in our 1-D model. Hence, it is reasonable to invoke the rotation in the midway of an acceleration during a time step. This Boris scheme uses two half-accelerations and one rotation in between. The overall motion of a particle in a time step is ([1] Birdsall),

1. First half-acceleration :

$$\begin{aligned} v_x(t') &= v_x(t-\frac{\Delta t}{2}) + \frac{q}{m}E_x(t)\frac{\Delta t}{2} \\ v_y(t') &= v_y(t-\frac{\Delta t}{2}) \end{aligned} \tag{140}$$

2. Rotation :

$$\begin{pmatrix} v_x(t") \\ v_y(t") \end{pmatrix} = \begin{pmatrix} \cos(\Omega\Delta t) & \sin(\Omega\Delta t) \\ -\sin(\Omega\Delta t) & \cos(\Omega\Delta t) \end{pmatrix} \begin{pmatrix} v_x(t') \\ v_y(t') \end{pmatrix} \tag{141}$$

3. Second half-acceleration :

$$\begin{aligned} v_x(t+\frac{\Delta t}{2}) &= v_x(t" - \frac{\Delta t}{2}) + \frac{q}{m}E_x(t)\frac{\Delta t}{2} \\ v_y(t+\frac{\Delta t}{2}) &= v_y(t" - \frac{\Delta t}{2}) \end{aligned} \tag{142}$$

where t' and t" are dummy variables. The angle of rotation, measured with respect to v_x-direction in counterclockwise sense, is $\Delta\theta = -\Omega\Delta t$. Note that the cyclotron frequeucy Ω carries the signs of q and B_o. The leap-frog method and Boris mover are implemented in the two subroutines MOVE and ACCEL respectively. It is possible to extent the Boris scheme, though more complicated, to include arbitrary values of θ_B ([1] Birdsall). In this case, all three velocities are taken into account.

The above-mentioned algorithm is essentially collisionless because it is possible for two particles to have the same (or nearly the same) positions at a time. As we discussed in Chapter 2, the collisional effect is unimportant on the time scales we are interested. It is thus justified to use a collisionless mover in the SEE simulation.

One complication arises at t=0 when the initial conditions, x(0) and $\mathbf{v}(0)$, are given at the same time. The main loop runs with x leading $\mathbf{v}$ by $\Delta t/2$. Hence at the start, $\mathbf{v}(0)$ is moved backward to $\mathbf{v}(-\Delta t/2)$ by running the Boris scheme backward. This is done by the subroutine SETV.

4.2.3 Field equation (Poisson's Equation)

In the electrostatic problem, we assume there is no induced electric field from the time varying magnetic field, that is $\nabla \times \mathbf{E} \approx 0$. Hence the electric field can be written as,

$$\mathbf{E} = -\nabla\phi \tag{143}$$

Combined with Gauss's law, it becomes the Poisson's Equation,

$$\nabla^2\phi = -\frac{\rho}{\epsilon_o} \tag{144}$$

It is an advantage to work with the scalar potential in Poisson's Equation. In our 1-D plasma simulation, it is simplified to,

$$\frac{\partial^2\phi}{\partial x^2} = -\frac{\rho}{\epsilon_o} \tag{145}$$

There are two approaches to solve this second order differential equation. One approach uses the center difference method and the other uses the fast Fourier Transform (FFT). In ES1 program, FFT is chosen because of its computational efficiency for large number of grid cells. An introduction to FFT will be given in Chapter 5 when we discuss the computation of power spectrum from a time series.

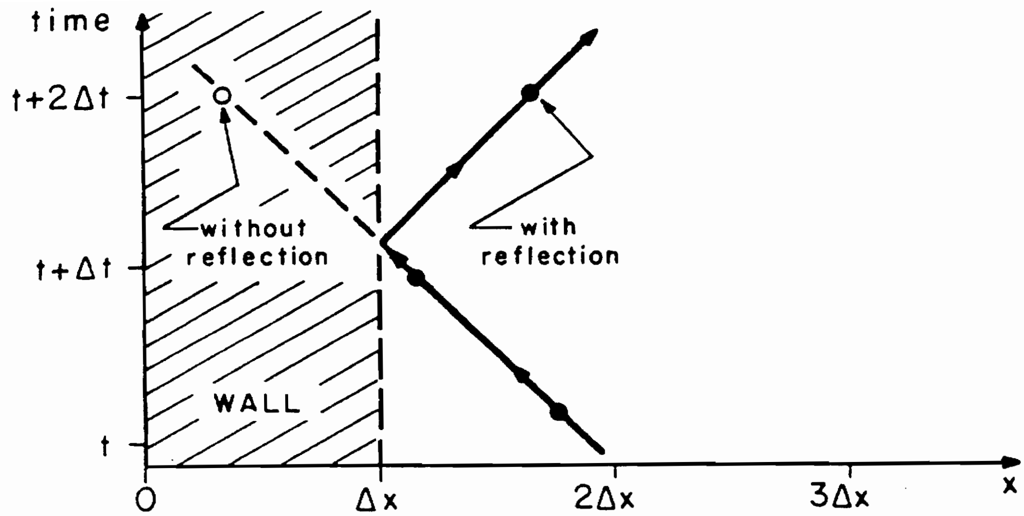


Figure 20: Reflective boundary condition used in ES1G (From [1] Birdsall).

A prerequisite to use FFT to solve the 1-D Poisson's Equation is that the boundary condition has to be periodic, that is the charge density and potential at the first grid cell are equal to those at the last grid cell. The plasma length under consideration is just a portion of an infinite plasma which has a spatial periodicity of l. When a particle reaches the two end points, it re-enters another boundary as if it continued to move without boundaries. The general algorithm of Poisson solver using FFT is,

$$\rho(x) \xrightarrow[FFT]{} \bar{\rho}(k) \xrightarrow[k^{-2}]{} \bar{\phi}(k) \xrightarrow[IFFT]{} \phi(x) \xrightarrow[\nabla\phi]{} E_x(x)$$

where $\bar{\rho}(k)$ and $\bar{\phi}(k)$ are FFT of $\rho(x)$ and $\phi(x)$ respectively. In the transform domain, the potential is easily obtained by,

$$\bar{\phi}(k) = \frac{\bar{\rho}(k)}{\epsilon_o k^2} \tag{146}$$

Then using inverse fast Fourier Transform (IFFT), we can get back the potential in x. By the center difference method, the electric field $E_x(x)$ is obtained.

With a grid system, the whole plasma length is essentially discretized or sampled in space. One has to make sure that the grid size is sufficiently small so that aliasing problems do not exist ([1] Birdsall).

In the ionosphere, the density increases linearly with height (x). We have to change the particle loading method, which will discuss in details in the next section, and the boundary conditions in our simulation program, called ES1G (G stands for density gradient). A clever way to apply the FFT method in the Poisson solver with a density inhomogeneity is discussed in [1] Birdsall. It is referred to as reflective boundary conditions, as shown in Figure 20. The method always keeps the first and last grid cells empty of particles by reflecting all the incident particles back, that is $x(t+\Delta t) = 2\Delta x - x(t)$ and $v(t+\Delta t) = -v(t)$ at the first grid cell. Then the electric fields vanish in these two extreme grid cells and the corresponding electric potentials must be modified by adding the solution $\phi = a + bx$ of the homogeneous equation $\partial^2\phi/\partial x = 0$ to the previous solution from FFT. The constant a is set to zero (reference potential only) and the constant b is evaluated by the boundary condition,

$$E_x(0) = -\left(\frac{\partial \phi_p}{\partial x}\right)_{x=0} - b = -\left(\frac{\partial \phi_p}{\partial x}\right)_{x=l} - b = E_x(l) = 0 \tag{147}$$

where ϕ_p is the particular solution obtained from FFT. Obviously, the two partial derivatives are equal because that is the assumption in FFT method. Hence, using the center difference method,

$$b = -\left(\frac{\partial \phi_p}{\partial x}\right)_{x=0} = \frac{\phi_p(L-\Delta x) - \phi_p(\Delta x)}{2\Delta x} \tag{148}$$

and the solution to the potential using the reflective boundary conditions is,

$$\phi(j) = \phi_p(j) + (j-1)b\Delta x \qquad \text{for} \;\; 1 \leq j \leq ng+1 \tag{149}$$

The Poisson solver is implemented in the subroutine FIELD, with reflective boundary conditions. Also, the reflective boundary conditions are included in the subroutine MOVE.

So far, we did not address the connection between grid and particle quantities. It is necessary to calculate the charge density on the discrete grid points from the continuous particle positions and to calculate the force at the particles from the fields on the grid points. These calculations are called weighting, which implies some form of interpolation among the grid points nearest the particle. It is desirable to use the same weighting in both density and force calculations in order to avoid a self-force which causes a particle accelerating itself ([1] Birdsall).

There are two types of weighting implemented in ES1 (as well as ES1G). They are:

- Zero-order weighting counts the number of particles within distance $\pm\Delta x/2$ (one cell width) about the j^{th} grid point and assigns that number to that point (Fig. a). The grid density is simply the collected number divided by the grid size. The common name for this weighting is nearest-grid-point (NGP). The same principle can be applied to force weighting. This method is simple to implement, however, it has an undesirable effect when a particle passes through a cell boundary. The charge density then jumps up and down at the two grid points associated with that cell boundary and becomes noisy both in time and space. This noise may be intolerable in many plasma problems.

- First-order weighting smooths the density and field fluctations, which is less noisy than the zero-order, but requires additional computation in accessing two grid points for each particle, twice per step. The weighting is actually a linear interpolation. The charge of a particle is proportionally divided to its nearest grid points, according to the distances between the particle and grid points. Mathematically, if a particle is situated in the j^{th} grid cell, then the charge assignments to the j^{th} and $(j+1)^{th}$ grid points are,

$$q_j = q\left(1 - \frac{x_i - X_j}{\Delta x}\right) = q\frac{X_{j+1} - x_i}{\Delta x} \tag{150}$$

$$q_{j+1} = q\frac{x_i - X_j}{\Delta x} \tag{151}$$

where X_j $(= (j-1)\Delta x)$ is the position of the j^{th} grid point. Note that the sum of q_j and q_{j+1} equals the particle charge q. This is sometimes called cloud-in-cell (CIC) model ([1] Birdsall) because the charged particles seem to be finite-size rigid clouds which may pass freely through each other. The field weighting operates in the same manner. That is,

$$E(x_i) = \frac{X_{j+1} - x_i}{\Delta x}E_j + \frac{x_i - X_j}{\Delta x}E_{j+1} \tag{152}$$

Higher-order weighting is possible to further reduce noise, but at the cost of more computation. It tends to reduce nonphysical effects introduced by the particle-in-cell method.

Before ending this section, we want to discuss how an external electric field $(E_o \cos\omega_o t)$ is added into the plasma during simulation. The first idea is to add the pump field amplitude to all calculated field amplitude in every time step. This is implemented in the original ES1 program. However, in the SEE problem, the external electric field may not spread over the whole plasma length. Thus, in ES1G program, the pump field is only effective over a fixed width ($espan$), centered at a predefined grid point ($ecenter$). They are depicted in Figure 21. We will give reasons why this is necessary in SEE simulation. As seen later, the finite span of the pump field is crucial to reproduce some of important features of SEE.

4.2.4 Particle position and velocity loading with density gradient

In classical mechanics, if one knows all positions and velocities of particles at a time, the subsequent motions can be exactly calculated. We have discussed the equation of motion and field solution. Now we will investigate how to load the initial positions and velocities of particles according to some distribution functions. This is done by the subroutine INIT. The density profile we are interested in SEE simulation is linear, as shown in Figure 21. Its shape is defined by the two positions, x_{min} and x_{max}, and the normalized density difference, $dfn = (n_{max} - n_{min})/n_o$. The density, n_o, at the center, $x_o = (x_{max} + x_{min})/2$, is derived from input parameters by,

$$n_o = N/lm \tag{153}$$

where N is the total number of particles of each species and $lm = x_{max} - x_{min}$ is the effective plasma length loaded with particles. The parameters x_{min} and x_{max} are normalized with respect to l when inputted.

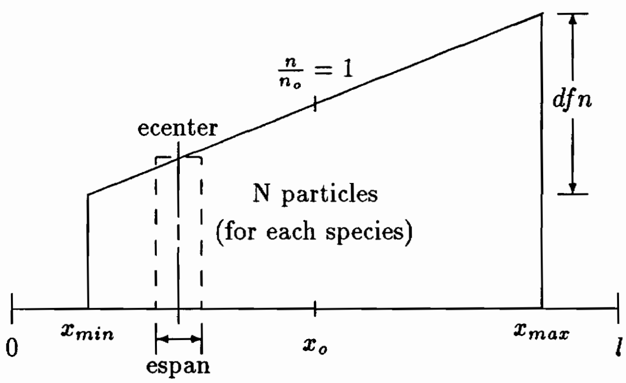


Figure 21: Linear density gradient used in ES1G program.

The distribution function $f_n(x)$ for the particle positions can be written as (using elementary slope-point form of a straight line),

$$\begin{aligned} f_n(x) &= \left(n_o - \frac{dfn \cdot n_o}{2}\right) + \frac{dfn \cdot n_o}{lm}(x - x_{min}) \\ &= n_o\left(1 + dfn\frac{x - x_o}{lm}\right) \qquad \text{for} \quad x_{min} \le x \le x_{max} \end{aligned} \tag{154}$$

By definition, the area under the curve of $f_n(x)$ is the number of particles. Hence, we have,

$$\int_{x_{min}}^{x_i} f_n(x)dx = i \qquad \text{for} \quad i = 1, 2, \cdots, N \tag{155}$$

Evaluating the integral, we get,

$$\frac{dfn \cdot n_o}{2lm}x_i^2 + \frac{N - dfn \cdot n_o x_o}{lm}x_i + \frac{\left(\frac{dfn \cdot n_o}{2}x_{max} - N\right)x_{min}}{lm} = i \tag{156}$$

or,

$$\frac{dfn \cdot n_o}{2}x_i^2 + (N - dfn \cdot n_o x_o)x_i + \left[\left(\frac{dfn \cdot n_o}{2}x_{max} - N\right)x_{min} - i \cdot lm\right] = 0 \tag{157}$$

This is a quadratic equation in x_i and its solution is,

$$x_i = \frac{\sqrt{B^2 - 4AC} - B}{2A} \qquad \text{for} \quad A \neq 0 \tag{158}$$

where A,B and C are the coefficients of x_i^2, x_i, and the constant term, respectively ($B \geq 0$). Note that the other solution with negative square root is rejected because the position x_i is always positive, and the product $4AC$ inside the square root is always positive. This formula also works for negative dfn. The procedure of position loading is simply to substitute every value of i from 1 to N, to get the corresponding particle position.

The quadratic formula fails to work when the parameter, dfn, is zero, that is all particles are uniformly distributed within the effective plasma length. For this case, another simpler formula is used.

$$x_i = \left(i - \frac{1}{2}\right)\frac{lm}{N} + x_{min} \tag{159}$$

Here, the discussion on position loading is completed. The loading procedure is repeated for all species. After position loading, the subroutine INIT invokes another subroutine SETRHO to calculate the background charge to neutralize the whole plasma. Also, the positions are normalized with respect to the grid size (dx) in the subroutine SETRHO.

The procedure of velocity loading is different because the distribution function $f_v(\mathbf{v})$ is more complicated. Usually, Boltzmann-Maxwellian or Gaussian distribution function is used to describe the velocity distribution in a equilibrium finite-temperature plasma. The distribution function f_v is of the form $exp(-v^2/2v_t^2)$, where v is the magnitude of the velocity $\mathbf{v}$ and v_t is the thermal velocity. One way to load the velocities is to use Gaussian random number generator. This is called a noisy start. Another way is to place the velocities according to the Gaussian distribution. This is referred as quiet start, which is used in both ES1 and ES1G programs ([1] Birdsall). Due to the complicated form of the Gaussian function, the velocity loading is done by numerical integration of f_v. We first define the cumulative distribution function F($\mathbf{v}$) as,

$$F(\mathbf{v}) = \frac{\int_0^v e^{-v^2/(2v_t^2)}dv}{\int_0^\infty e^{-v^2/(2v_t^2)}dv} \tag{160}$$

It has a range between 0 and 1. Then F($\mathbf{v}$) is set equal to a set of uniformly distributed numbers varying from 0 to 1. That is,

$$F(\mathbf{v}) = \frac{1}{\mathrm{N}}, \frac{2}{\mathrm{N}}, \cdots, 1 \tag{161}$$

A DO-loop is used to find out all velocities of particles. Then the velocity v is equally assigned to all three directions. The last step in velocity loading is to decorrelate the phase space by randomly exchanging the indices of positions and velocities in pairs.

In plasma simulation, we do not directly work with all the four basic quantities: charge (q), mass (m), time (t) and length (l). These quantities cannot directly tell the major characteristics of a plasma. The plasma frequency (ω_p), cyclotron frequency (Ω), thermal velocity (v_t) and charge-to-mass ratio (q/m) are more relevant in the definition of a plasma. Therefore, the latter set of parameters are chosen as input to derive the basic quantities. The charge is calculated from the plasma frequency (see (11)) as follows.

$$Q = \frac{\epsilon_o \omega_p^2}{q/m} \left(\frac{l}{N} \right) \tag{162}$$

In one-dimensional model, the definition of "charge" must be modified to be compatible with the dimension of the linear density (N/l). It is actually the charge per unit length (Q). Similarly, the mass per unit length can be calculated from the q/m ratio. The unit time is determined by the plasma frequency and the unit length is defined by the effective length of the plasma. The magnetic field (B_o) is calculated from the cyclotron frequency formula (see (17)).

$$B_o = \frac{\Omega}{q/m} \tag{163}$$

Note that the q/m ratio carries the sign of charge, whereas the cyclotron frequency carries the signs of charge and magnetic field.

4.2.5 Precautions about numerical accuracy

There are a number of points worth for our attention when we set up a plasma model to simulate. The first one is the requirement of time step ([1] Birdsall).

$$\omega_{pe} \Delta t \leq 0.2 \tag{164}$$

It is a common practice to set the time step to the order of $0.2/\omega_{pe}$. In inhomogeneous plasma, the plasma frequency depends on the position in the plasma. One should use the plasma frequency in the region of the greatest interest to determine the time step.

The next accuracy requirement is about the limitation of grid size imposed by Debye length. We see that in the first-order weighting, the particle clouds in a spatial grid lose their short-range interactions and pass smoothly through one another with relatively small noise. In addition, it is shown that for cloud size less than or equal to a Debye length ($\Delta x \leq \lambda_D$), longitudinal waves and Debye shielding are nearly the same as for a laboratory plasma ([1] Birdsall). In an inhomogeneous plasma with steep density gradient and $\theta_B = 0$, significant diffusion across the background magnetic field may result if a grid cell consists of many Debye lengths.

Our numerical model has to fulfill the criterions for an ionized gas to be a plasma (section 3.3). The first requirement is that the Debye length is small compared with the other physical dimensions of interest. Since a Debye length is at least of the order of a grid cell, this gives an idea of how small a grid cell should be. However, as we mentioned in section 4.2.3, the grid size has to be small enough to avoid aliasing. Mathematically, from sampling theory,

$$k_{max} \Delta x \leq \pi \tag{165}$$

where k_{max} is the maximum wave number that occurs in the simulation. Also, since the plasma length is finite, it has to be long enough to allow events with the longest wavelength or the smallest wave number (k_{min}) to occur in the simulation. In other words,

$$k_{min} \geq \frac{2\pi}{l} \tag{166}$$

The second requirement is that the number of particles is large enough in a Debye sphere. In 1-D plasma, the Debye sphere is equivalent to a Debye length. A common practice is to set a minimum number of particles per grid cell greater than 20, if the Debye length is approximately equal to the grid size. Note that even though the grid size is of the order of a Debye length, it does not necessarily mean there are enough particles per grid cell, and vice versa. One has to check both requirements.

As seen from (163), the charge per unit length (Q) depends on the plasma frequency and the q/m ratio. Therefore, the number of significant digits of ω_p and q/m is important to maintain charge neutrality of a plasma. One should try to use at least five significant figures to avoid unnecessary numerical errors due to unbalanced charges between particle species. The final caution is about the settings of x_{min} and x_{max}. Do not attempt to set x_{min} to zero and x_{max} to 1. It is because after position loading, the loaded particles in the first and last grid cells are reflected in the subroutine SETRHO. Then two glitches in the density profile result.

4.2.6 Diagnostics

In ES1, there are several built-in diagnostics to verify whether the result is reasonable or not. For particles, they are plots of phase space, velocity space and velocity distribution functions. The plotting interval can be arbitrarily set. For grid quantities, one may plot charge density, potential and electric field versus position. Moreover, at the end of simulation, the history of a run is summarized in plots of field energy, particle kinetic energy, particle drift energy, particle thermal energy, total energy and mode energy versus time. For more details, one may refer to [1] Birdsall.

A number of diagnostic plots is added in ES1G program to facilitate the SEE investigation. They are:

- Density plots for each species at some regular intervals during run.
- History plots for traces or trajectories of test particles.
- Animation of the density profiles for each species.

A sample run can be found in Appendix (C). This is, in fact, a case study of the SEE simulation which will discuss in the next chapter.

5 Numerical simulation of SEE in the ionosphere

We devote this chapter to the most important work of this thesis. The numerical simulation of SEE presented here is believed to be the first plasma simulation of this phenomena. The main goal of this simulation study is to supplement the experimental observations and future theoretical work. In many aspects, plasma simulation can provide more insights of dependence of the SEE spectrum on various physical parameters (e.g. thermal velocities, etc.) than the experiments.

In this thesis, only the interactions at the upper hybrid layer are investigated. The region near the reflection layer will be left for future study. We use a 1-D 3V PIC electrostatic code (ES1G) to run all simulations mainly due to its simplicity and computational efficiency. We also believe that an 1-D model is sufficient to reproduce some of the SEE features such as the DM and BUM, which are likely to be generated in the upper hybrid layer. Two species are used because the dominant ion species is O^+. We assume that an electrostatic upper hybrid wave which is mode-converted from the O-mode pump is the source of SEE generation at the upper hybrid layer. This pump field is only effective over a small region around the upper hybrid point. The magnetic field is set perpendicular to the density gradient, and hence only two velocities (v_x and v_y) are effective in the simulation.

This chapter is organized as follows. First, the calculation of the simulated power spectrum using the fast Fourier Transform (FFT) will be briefly reviewed. Next, we will discuss the simulation setup which is quite important to get valid and meaningful simulation results. Then we go on to display various simulation results from which we have new results about the DM, UM, BUM and quenching mechanisms of DM. All of them will be summarized in our discussion section.

5.1 Calculation of power spectrum

A time series of the electric field[1] is the major output of the plasma simulation for the SEE study. Since SEE features are classified in the frequency domain, some type of Fourier Transform of the electric field time series is needed to obtain power spectra for direct comparison with the experimental data. The fast Fourier Transform (FFT) technique is used for such a purpose. This is historically called the periodogram method for power spectral estimation. In the following, it will be briefly reviewed.

An energy-bounded signal $h(t)$ and its Fourier Transform H(f) are related by the Fourier Trans-

[1] It is contained in the file ES1G.DAT. The file format is discussed in Appendix (C).

form pair.

$$h(t) = \int_{-\infty}^{\infty} H(f)\, e^{-i2\pi f t}\, dt$$
$$H(f) = \int_{-\infty}^{\infty} h(t)\, e^{i2\pi f t}\, df \tag{167}$$

Sometimes, the first integral is called the inverse Fourier Transform of $H(f)$. Under the Fourier Transform, periodic signals produce delta functions at their harmonic frequencies, while aperiodic signals produce a continuous spectrum. Detailed discussion of Fourier Transform can be found in many engineering mathematics text[2].

In our plasma simulation, we are interested in the discretized version of Fourier Transform since we must compute spectra of the electric field time series from the plasma simulation code. If we have N consecutive sampled values[3] of $h(t)$,

$$h_k = h(t_k) \qquad \text{with } \; t_k = k\Delta t, \quad k = 0, 1, 2, \cdots, N-1 \tag{168}$$

where Δt is the sampling time interval of $h(t)$. The discrete Fourier Transform of h_k exists at the discrete values of frequency,

$$f_n = \frac{n}{N\Delta t}, \qquad n = -\frac{N}{2}, \cdots, \frac{N}{2} \tag{169}$$

The Fourier integral is approximated by,

$$H(f_n) \approx \Delta t \left(\sum_{k=0}^{N-1} h_k e^{i2\pi f_n t} \right) \tag{170}$$

The sum inside the bracket is called the discrete Fourier Transform of h_k and is denoted as H_n. Hence,

$$H(f_n) \approx \Delta t H_n \tag{171}$$

The discrete Fourier Transform maps N complex numbers (h_k) into N complex numbers (H_n). For a real signal (which is the case we are considering), its discrete Fourier Transform is symmetric about zero frequency, that is $H_n = H_{-n}$. Similarly, one can obtain a formula for the discrete inverse Fourier Transform ([13] Press).

In a sampled data system, the Nyquist criterion specifies the minimum sampling frequency for a given signal $h(t)$ or equivalently, the maximum output frequency from the discrete Fourier Transform. On the other hand, The smallest frequency separation is $1/N\Delta t$ (refer to (169)). The frequency

[2]For example, Fourier series and boundary value problems (4th edition), by R.V. Churchill and J.W. Brown (McGraw Hill).

[3]For the simplest case, we assume uniform sampling.

resolution increases with N. As the number of input data increases, the output frequency spectrum becomes closer to the continuous spectrum. Thus we may be interested in how much computation it requires for a N-point discrete Fourier Transform. Define a complex number W as,

$$W = e^{i2\pi/N} \tag{172}$$

Then the discrete Fourier Transform can be rewritten as,

$$H_n = \sum_{k=0}^{N-1} W^{nk} h_k \tag{173}$$

Obviously, it takes N^2 complex multiplications and N summations to complete the transform. So, the discrete Fourier Transform appears to be an $O(N^2)$ process. For many years, people believed that was the answer. But, in 1940s to 60s, a number of people noticed that there were some redundant multiplications in the discrete Fourier Transform. Since the factor W^j is cyclic for $|j| > N$, the number of computations of the product $W^{nk}h_k$ can be reduced dramatically if we use a clever algorithm to remove all the redundant multiplications. This is called the fast Fourier Transform[4] or FFT. Derivation of FFT algorithm is quite involved. Therefore we leave it to many good texts on digital signal processing (e.g. [13] Press). The computational complexity of FFT is reduced to $O(N \log_2 N)$ which is a very significant improvement for large N. In most of the cases, FFT requires N to be an integer power of two for an efficient binary manipulation of its algorithm.

An example FFT subroutine[5] called FOUR1 can be found in [13] Press. The FOUR1 subroutine is incorporated into a program we developed to calculate power spectra[6] called PSD. The power spectrum is formed by,

$$P_n = |H_n|^2 = H_n H_n^* \tag{174}$$

at each output frequency of the FFT. The listing of PSD program is included in Appendix (B).

The PSD program is equipped with a number of features to facilitate our efficient study of the SEE simulation. The input electric field time series can be selected as a part of the whole series outputed from ES1G program. This allows us to skip transient effects[7] and to search for the characteristics of SEE spectra within different portions of the time series. Also, the PSD program is built in zero patching for a selected part of the electric field time series which may not be exactly

[4] In fact, there exist many algorithms and derivations of FFT. Interested readers may refer to [13] Press.

[5] It is written in FORTRAN. The input and output arrays are also specified in [13] Press.

[6] For an aperiodic signal, its total energy is finite. Fourier Transform of such a signal gives an energy spectrum. On the other hand, power spectrum or power spectral density is formally related to periodic signals. However, under many circumstances, the power spectrum is apparently equivalent to the energy spectrum in literatures, as well as in this thesis.

[7] Recall that the experimental SEE spectrum is taken under steady state conditions ([11] Leyser).

an integer power of two. However, one should pick the number of desired data points close to an integer power of two to avoid unnecessarily high noise level. Two commonly used numbers of data points are 8192 and 16384 for the calculation of power spectra displayed in section 5.3. Also, three types of filtering methods are implemented in the PSD program to reduce noise effect on the desired spectra. They are:

- Smoothing the input electric field time series by weighted three-point average[8].
- Smoothing the computed power spectrum by weighted three-point average.
- Doubling the number of input electric field time series by linear interpolation[9].

Besides filtering, the PSD program provides four FFT windowing[10] selections. They are square, parabolic, raised cosine, and bilinear windows. It enables users to trade off between leakage effect (sidelobe levels) and frequency resolution (beamwidth of the main lobe). Usually, the square window with power spectrum averaging is used in the calculation of the displayed power spectra in this thesis. Magnification of part of the power spectrum is also available in the PSD program. A sample run of PSD program can be found in Appendix (C).

We mentioned that the Poisson solver in ES1G program uses FFT to solve for the potential from a charge density. The computational complexity for FFT is $O(10240)$ if 1024 grid cells are used. If one tries to directly calculate the potential other than using FFT, the number of computations will be $O(1024 \times 1024)$, which is much more than using FFT. More details can be found in [1] Birdsall.

Besides the FFT method, there exist some other numerical methods which can have a better frequency resolution, if properly implemented, without any increase in the number of input data points. The maximum entropy method is such a method, which has a very cute property of being able to fit sharp spectral features. One can look up FORTRAN subroutines of this method in [13] Press. But one should be very careful when using the maximum entropy method, because the computed power spectrum can deviate a lot from the real one in some cases. The maximum entropy method may be useful for studying the narrow DP feature.

[8] The weights of the center point and its two adjacent points are 1-2-1.

[9] Note that the frequency resolution does not increases in this case because the sampling interval decreases by half although the number of data points becomes twice.

[10] Sometimes, we can get frequency components other than those expected in the input signal from the FOUR1 subroutine, if the first and last input data are nonzero. This is called aperture or leakage effect of a square window. Thus, various window functions are proposed to minimize the leakage effect, at the expense of reducing resolutions. A more thorough discussion on FFT windowing can be found in [13] Press.

5.2 Simulation setup

In order to reproduce some of the SEE features, a correct simulation setup is essential. In this section, we want to discuss the appropriate range of parameters used in ES1G program. As we have already discussed in section 4.2.4, plasma parameters such as the number of particles of each species (N), effective system length (lm), plasma frequency (ω_p), cyclotron frequencies (Ω), charge-to-mass ratio (qm), and the thermal velocities (v_t) are more relevant in the definition of a plasma and are chosen as input to specify the plasma in SEE simulations. Although these parameters can be arbitrarily set[11], their settings are still bounded by some computational limitation and convenience. N is set to 40000 for a reasonable simulation time without too many nonphysical phenomena (e.g. noise). For computational simplicity, the parameters qm_e and v_{te} for electrons (the first species) are set to -1 and 1, respectively. The setting of the nominal electron plasma frequency ω_{peo}[12] depends on the value of the electron cyclotron frequency Ω_e because their ratio, ω_{peo}/Ω_e, is approximately determined by the harmonic number of electron cyclotron frequency which we are interested in investigating. For the third and fourth harmonics, Ω_e[13] is set to -0.3 so that the plasma frequencies are around unity. Then, according to the ionospheric models tabulated in section 2.3 (Table 1), ω_{peo} can be determined. For example, if we assume $f_{ce} = 1.36MHz$ in the ionosphere, for Region I, the corresponding nominal plasma frequency for the simulation model is $\omega_{peo} = 0.3 \times 3.47/1.36 = 0.765$. When ω_{peo} is known, for a given pump frequency ω_o within the range of the three model regions, one can use a program called UH (stands for upper hybrid point) to calculate the upper hybrid position (in terms of the grid position) within the plasma system and the corresponding lower hybrid frequency. Then one can set the center position (ecenter) of the pump field equal to the upper hybrid point for the ES1G simulation. The UH program uses the formula (59) given in Chapter 3. Its program listing is shown in Appendix (E).

[11] In plasma simulation, the ratios between parameters are more important than the individual parameter value. The absolute value of a parameter is meaningful only if it is defined with a specified unit. Usually, we do not define the physical units (e.g. meters, seconds, etc.) for parameters used in plasma simulation because in most cases, it is impossible to simulate a laboratory or ionospheric plasma with the same number of particles inside a Debye sphere (or a Debye Length). Using the Region II of our ionospheric model in Figure 4 as an example, if we want to simulate with the same number of particles as in Region II, it requires a huge number of particles equal to $(3.5 \times 10^{11}\, m^{-3})^{1/3} \times (45\, Km) = 0.317 \times 10^9$, which is impossible to simulate using any existing supercomputers. Only when the number of particles inside a Debye sphere is equal, there exists a one-to-one correspondence of units between the two systems. Hence, all the four basic units between the ionospheric plasma and the simulated plasma in general cannot be one-to-one related as in the MKS and Gaussian systems.

[12] This is the plasma frequency at the center of the effective plasma length x_o.

[13] Recall that the minus sign is due to the charge of electron.

Table 3: Parameter sets used in ES1G simulation for the three defined regions.

Region	ω_{peo}	Δt	dfn	lll	$espan$	$ngavg$
I	0.765	0.25	1.333	175.6	31	7
II	1.172	0.2	1.314	75	31	7
III	1.42	0.15	0.330	51.1	121	27

The system length (l), the number of grid cells (ng), the pump field amplitude (e_o), the effective span of the pump ($espan$), the number of grid cells for averaging the output electric field (to the file ES1G.DAT) centered at an input parameter ($ngpsd$), and the mass ratio of ions to electrons (K) have all been varied in a number of simulation runs. Then they are set as follows. Using the criteria outlined at the end of Chapter 4, we find that 1024 grid cells are sufficient for the SEE simulation. The values of the system length in terms of 2π ($lll = l/2\pi$) are shown in Table 3. Note that the effective system length is always less than the whole system length because x_{min} and x_{max} are usually set to 0.05 and 0.95[14], respectively to avoid effects by the artificial reflective boundary conditions implemented in ES1G program. Once the effective system length and the nominal plasma frequency are specified, the charge and mass of that species can be determined by (132). The electron charge and mass are the same for all three model regions. Also, from ω_{peo}, the time step Δt can be estimated using a formula given in section 4.2.5.

Table 3 summarizes some important simulation parameters for all three regions. The values of dfn are obtained from Table 1. For the setting of $espan$ and $ngavg$[15], an upper bound for $espan$ is that the upper hybrid wave should not extend beyond the plasma reflection point because the electric field is essentially parallel to the magnetic field near the reflection layer. Also, $ngavg$ should be smaller than $espan$[16]. A square spatial window[17] is used to distribute the pump field within the spatial region defined by $ecenter$ and $espan$. Typically, we use $espan = 31$ and $ngavg = 7$ for the first two model regions. Most likely, $espan$ and $ngavg$ depend on the density gradient. Their values

[14] When one uses the UH program, the normalized effective system length is $x_{max} - x_{min}$. Its default value is 0.9.

[15] As far as the author knows, there is no published literature providing the effective spatial widths of interaction and the spatial shapes of the electric field distribution for the forward and backward mode conversions.

[16] The parameter $espan$ corresponds to the spatial width of which the upper hybrid wave exists in the forward mode conversion, while $ngavg$ is the spatial width of which the backward mode conversion occurs.

[17] In the very early simulations, the pump field is spread over the whole plasma length. SEE features have never been observed because the widely spread pump field significantly disturbs the growth and propagation of some plasma waves which are crucial to SEE generation. Also, averaging the output electric field over a small number of grid cells is introduced because we discovered that the power spectra at two adjacent grid positions are very different. The averaging allows us to obtain a more representative spectrum of the waves around an observation point.

for Region III are set accordingly.

The settings of the pump field amplitude (e_o) and the mass ratio (K) are more difficult because we need to ensure that the interactions inside the plasma fall into the class of weak tubulence (see Figure 13) in order to generate SEE features. The threshold of parametric decay instability involving lower hybrid waves depends on the ion temperature and the mass ratio. The ion temperature is set to one fourth of the electron temperature as in the situation of the ionosphere. A small mass ratio favors a lower threshold and a faster growth rate of parametric instability, as well as a wide frequency separation between the pump and the DM. On the other hand, a simulation with high mass ratio may take a long time to run[18] before it can reach steady state. Moreover, too small a mass ratio may change the dominant physical process amongst the pump wave, electrons and ions. The mass ratio is set to 49 and the pump field amplitude is set to 0.15 in almost all simulation runs so that the PDI threshold is well exceeded and the whole plasma is not significantly perturbed for these values.

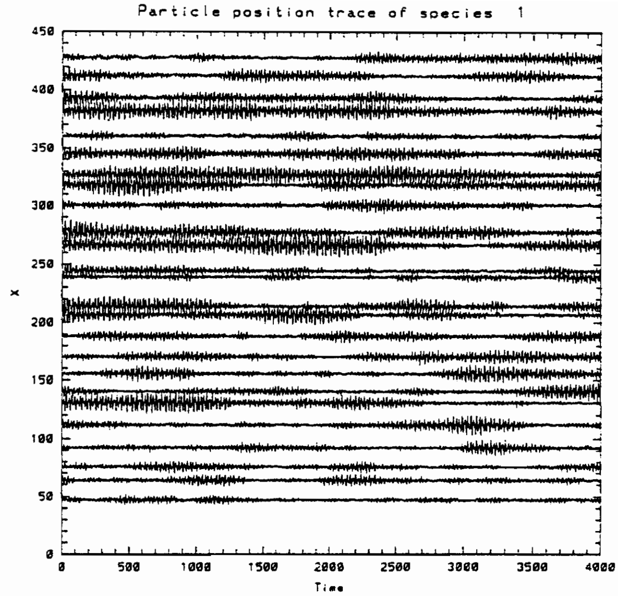
The setting of the number of time steps (nt) is determined by the following considerations. For a complete development of parametric instability involving the lower hybrid wave, simulation time has to be long enough to include several periods of the ion motion. The ion motion is crucial to the PDI as discussed in section 3.8. Moreover, the number of time steps should be compatible with the required number of data points for the calculation of power spectrum. Inital transient effects should be taken into account. In our SEE simulation, 20000 time steps are often used so that one can use 2^{14} (=16384) electric field data or less for a power spectral calculation. The run-time of ES1G program is about 9500 seconds using the IBM R6000 workstation[19]. Approximately six ion gyration cycles are contained in the whole simulation.

Figure 22 shows some simulation outputs for a case without pump field (noise case). It gives us some idea of how the noise spectrum looks like with Region II parameters. We can compare it with the driven SEE simulations to determine whether there is power enhancement at the interested frequencies. The first two pictures are electron and ion trajectories for 25 test particles[20] of each species. They are equally spaced within two normalized positions xo_{min} and xo_{max} (set to 0.1 and 0.9 in Figure 22) when the particles are initially loaded. As expected, the cyclotron frequency of electrons is much higher (49 times in Figure 22) than that of ions, and the Larmor radius of electrons

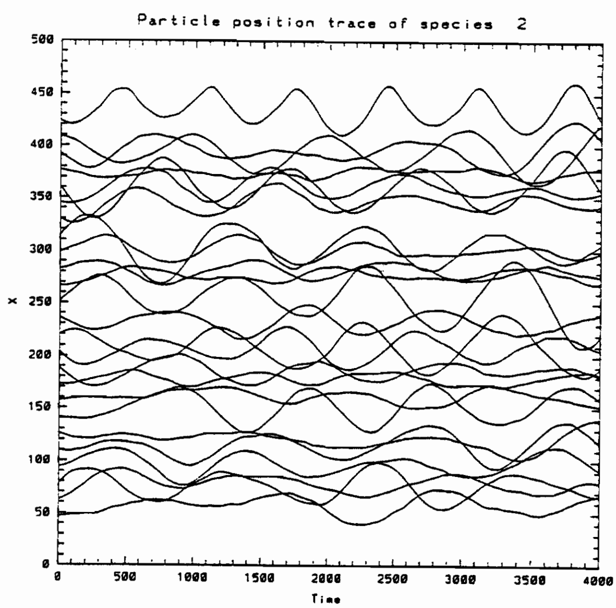
[18] A long simulation time may be undesirable because the numerical error accumlates as the simulation time goes on.

[19] It is a RISC based UNIX workstation introduced about two years ago by IBM. Its landmark performance is approximately 50MIPS. In our machine, the total capacity of random access memory is 96MB and the hard drive capacity is 2.0GB.

[20] It can be set to any number below 50 by changing the parameter npt.

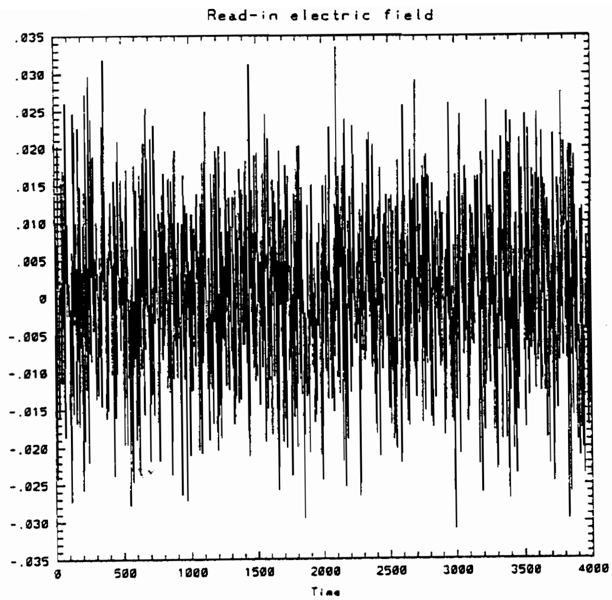


(a) Electron trajectories.

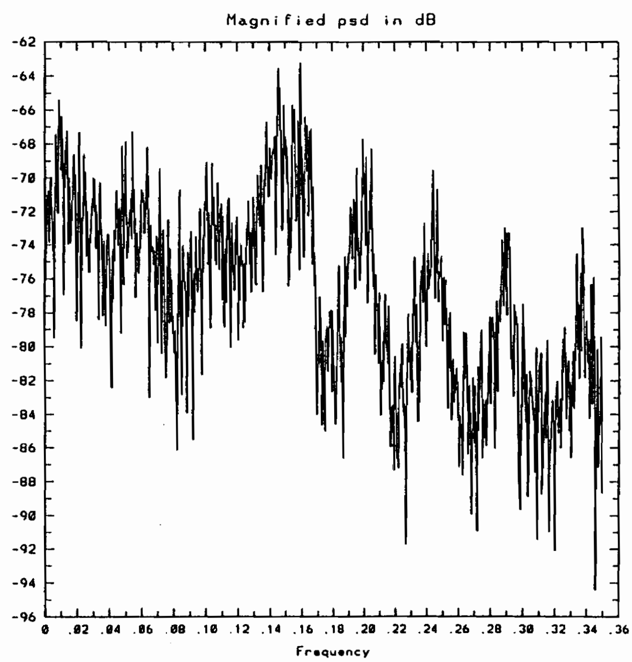


(b) Ion trajectories.

Figure 22: ES1G simulation without pump field (noise case).



(c) Electric field time series (ngpsd=154).



(d) Noise power spectrum from zero frequency to 0.35. ($t = 600$ to $t = 3800$)

Figure 22 (continued).

is much smaller than that of ions. Note that the amplitude of the electron gyration seems to be modulated in Figure 22(a). This is because the trajectories are sampled at a time interval much larger than the time step. Only 500 points are plotted for each trajectory. The variation of amplitude among trajectories is due to the velocity distribution. Figure 22(c) depicts the electric field time series outputted from ES1G program with *ngpsd* = 154, and Figure 22(d) is its corresponding power spectrum. The time series is truncated from $t = 600$ to $t = 3800$[21], as shown in the caption of the figure. Note that the power spectrum is plotted in dB scale. The noise floor is around -70dB for frequencies near the upper hybrid frequency (0.14) of the observation point. For frequencies much larger than those in the figure, the noise floor drops down to -110dB. The systematic variation in the noise spectrum is caused by thermal fluctuations in the plasma which trigger electron Bernstein modes ([1] Birdsall). It is evident that in frequency regions above the upper hybrid frequency, the thermal energy is localized in narrow bandwidths above the electron cyclotron harmonics, and in regions below the upper hybrid frequency, the thermal energy spreads out over the entire harmonic bandwidth.

A detailed setup of the reference case ($\omega_o = 0.88$) can be found in Appendix (C) and a list of all simulated cases discussed in this thesis, together with other less important cases, is tabulated in Appendix (D).

5.3 Simulation results

Using the simulation setup we discussed above, a large number of simulation runs have been performed to study the SEE phenomena. We are particularly interested in the cases which the pump frequencies are near electron cyclotron harmonics because from the experiments, most of the SEE features occur in these frequency ranges. The simulation results are presented in four groups: f_o slightly below nf_{ce}, $f_o = nf_{ce}$, f_o slightly above nf_{ce}, and cases other than the above, where $n = 2, 3, 4, 5$. Most of simulation parameters are same as mentioned in the last section, except those parameters unique to each case such as ω_o, *ecenter* and *ngpsd*. If different settings are used, they

[21] It is understood that the time unit has a value proportional to ω_{pe}^{-1}. But, in an inhomogeneous plasma, ω_{pe} does not remain constant. It is irrelevant to display the simulation time as $\omega_{peo}t = \cdots$ which can be found in most of published plasma simulation literatures because the local plasma frequency in general is not equal to ω_{peo}. If instead of ω_{peo}, the local plasma frequency ω_{pe} is used, then $\omega_{pe}t$ approximately represents the number of plasma oscillations at the pump region. Since the time step is fixed for all three model regions, it is difficult to compare the time from case to case with two different values of ω_{pe}. Moreover, in almost all circumstances, the time is used as a label to indicate which portion of the time series is taken for the calculation of power spectrum. Therefore, in this thesis, we refer to the time directly by its numerical value.

will be specified in the caption of figures and/or in the paragraphs.

5.3.1 Pump frequencies slightly below nf_{ce}

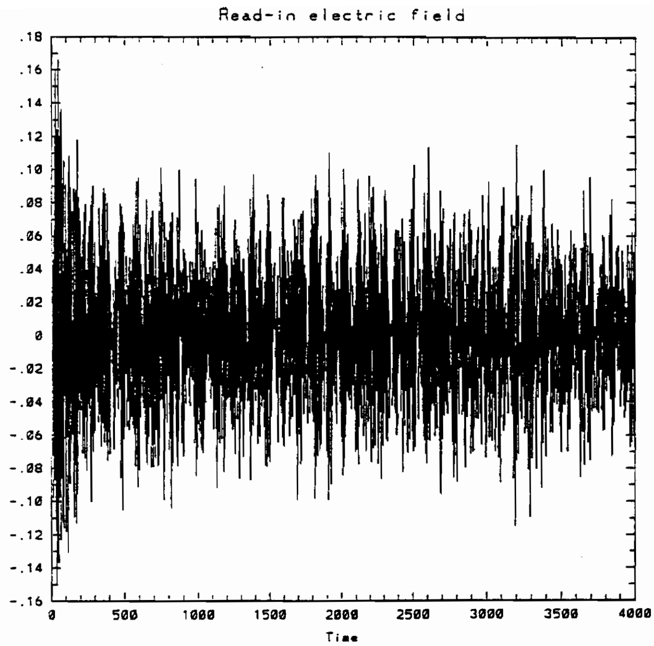
Third harmonic

We begin with pump frequencies slightly below the third harmonic. Figure 23 depicts some of the most important results of the reference case $\omega_o = 0.88$[22]. The pump and observation regions are both near the left edge of the plasma. Figure 23(b) and 23(c) show the magnified portions of the computed spectra around the pump frequency (f_c stands for center frequency which is usually set equal to the pump frequency.) for different portions of the electric field time series shown in Figure 23(a). The transient interaction inside the plasma is settled after $t > 300$. The lower sideband which is approximately 0.01 away from the pump is believed to correspond to the DM feature. Note the sharp cutoff at a frequency of 0.007 below the pump. The lower hybrid frequency of this simulation setup is about 0.0064 (calculated using the UH program). This provides another strong evidence to believe that the lower sideband corresponds to the DM. The spectral width of the pump in Figure 23(b) is narrower than Figure 23(c) because the frequency resolution of Figure 23(b) is twice of Figure 23(c). But, the bandwidth of the DM seems unchanged in both spectra and is about 0.006. The power spectrum below -60dB is irrelevant to the investigation of SEE generation because it is too weak and too irregular.

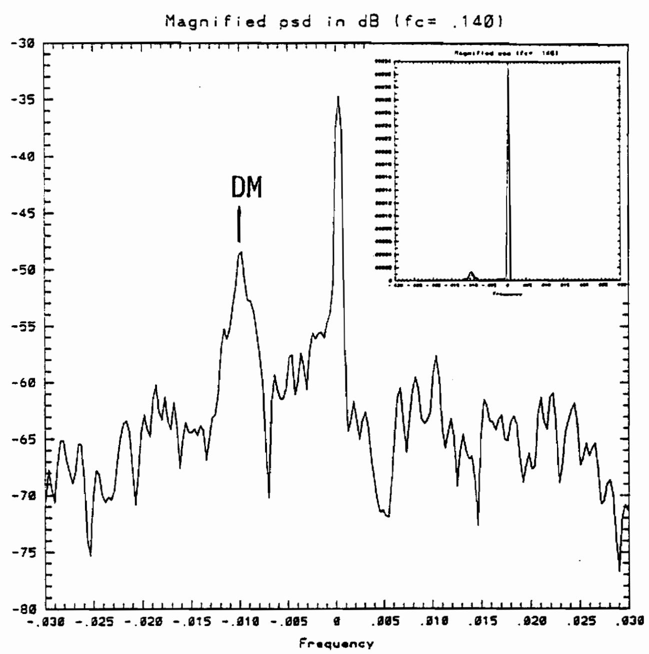
A strong low-frequency spike appears in the insert of Figure 23(c). Its frequency is approximately equal to the offset frequency of the DM from the pump. This frequency matching partially verifies parametric decay instability which occurs around the observation point and generates the DM feature. A complete confirmation of PDI requires wave vector matching. Usually, the wave vector matching is more difficult to verify in plasma simulation with inhomogeneity because of a number of reasons. First, the frequencies of the three waves involved in the PDI have to be known in advance. The electric field is then correlated with these three frequencies. FFTs in spatial dimension at a particular time instant may be used to measure the wave numbers. However, in an inhomogeneous plasma, the wavelengths vary spatially. Hence, the wave number matching is in general not as clear as the frequency matching. More details about the wave number matching can be found in [14] Lin 1982.

Comparing Figure 23(d) with the noise spectrum in Figure 22(d), we immediately observe power

[22] All the frequency measures in this thesis are with respect to the electron cyclotron frequency ($|\Omega_e| = 0.3$) which is constant in all simulation cases. For our presentation simplicity, all frequencies are labelled by their numerical values and are understood that they are meant relative to the electron cyclotron frequency.

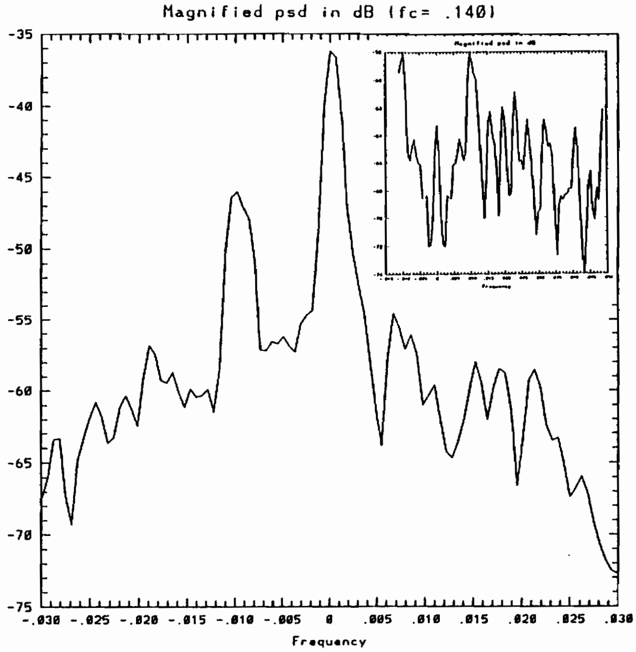


(a) Electric field time series.

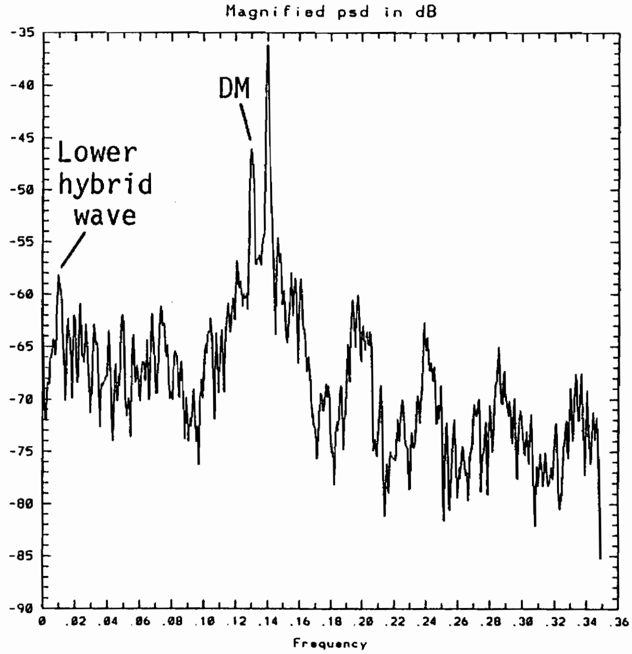


(b) Power spectrum for the time series from $t = 600$ to $t = 3800$ (16000 points). The inserted spectrum is in linear scale.

Figure 23: Simulation results at $\omega_o = 0.88$, with *ecenter* = *ngpsd* = 154 and Region II parameters.



(c) Power spectrum for the time series from $t = 1000$ to $t = 2600$ (8000 points). The inserted spectrum is a magnified spectrum around the zero frequency. Note that the strong frequency spike at 0.01 corresponds to the frequency offset of the DM from the pump.



(d) Comprehensive spectrum from zero frequency to 0.35. ($t = 1000$ to $t = 2600$)

Figure 23 (continued).

enhancement at several frequencies besides the pump and the DM. In general, the power over the whole spectrum in Figure 23(d) is enhanced by 5 to 6dB. Scanning from zero frequency, the first enhancement is at 0.01, which is thought to be a signature of PDI involving lower hybrid waves. Then the second enhancement is around the pump and DM. The next is at frequencies slightly above $4f_{ce}$ (0.191). It is also true for the higher cyclotron harmonic frequencies, but with decreasing power enhancement. Most probably, the electron Bernstein modes are triggered because most of the power enhancement is localized around the upper hybrid point, the third and fourth electron cyclotron harmonics, where their bandwidths are larger than the higher harmonics.

Some of the diagnostic printouts of this case ($\omega_o = 0.88$) are collected in Appendix (C) as an example. One can find plots of density profiles, charge density, potential, electric field and phase spaces at $t = 0$ and $t = 1400$, as well as the simulation history. Note that at $t = 1400$, a narrow peak appears in both electron and ion density profiles (Figure C-2(a) and (b)). The first species is electrons and the second is ions. In fact, with reference to the subsequently plots or the animation, this peak continues to travel towards the right hand side. The density profiles are not significantly perturbed. This peak is suspected to be a soliton because its shape does not change much until it ends its solitary propagation near the end of the plasma. This is common whenever the DM is observed in the power spectrum. However, when one views its propagation using an animation (ES1G.HDF), a number of similar stuctures can occur at the same time along the plasma, especially after one third of the simulation. The scenario resembles that a rod is dumped into a water tank and it generates large ripples propagating away from the excitation. Of course, the water tank does not have density gradient. When the phase spaces of electrons and ions at $t = 1400$ (Figure C-2(f) and (g)) are compared, it is evident that the tubulence is primarily due to ions, and electron motion is influenced by the ion motion. The last two history plots (Figure C-3(i) and (j)) show the trajectories of electrons and ions. Comparing with Figure 22(a), the electrons are heated around the pump region ($x = (ecenter/ng) \times l = 70.87$[23]). But, the power enhancement of electrons above the pump region (e.g. $x = 150$) seems to be delayed by a time period proportional to its distance from the pump. This is best explained with the ion trajectories. Note that there are some waves superimposed on the ion gyromotion trajectories. They emanate from the pump region and propagate apparently towards positive x-direction with decreasing velocities and amplitudes. An estimate of the "average"

[23] Similar to the time and frequency, the displayed lengths in this chapter are meant with respect to some convenient length parameters such as the Debye Length (λ_D) or the system length (l). However, the Debye Length varies with positions in the plasma. The system length seems to be a more reasonable choice of the base reference unit for length. In nearly all situations, the length is used for labelling. Hence, it is displayed by its numerical value and is understood that it is with respect to the system length.

propagation velocity of these waves is $(400-70)/4000 = 0.0925$[24]. This can account for the electron and ion density fluctuations at $t = 1400$. If one carefully reckons the oscillation frequency of these parasitic waves, it is approximately equal to the low-frequency spike observed in the Figure 23(c). Putting all the facts together, we can conclude that these waves are lower hybrid waves and their existence is necessary for the generation of the DM feature.

Figure 24(a) shows the power spectrum when the pump frequency is shifted to $\omega_o = 0.86$. The frequency offset of the DM peak from the pump increases by 0.002 when compared with Figure 23(c). Note that a weak frequency peak which appears at about a frequency of 0.02 below the pump may be the second DM. It also shows up in Figure 23(c). Another weak sideband appears at a frequency of 0.011 above the pump in Figure 25(a). It is suspected to be the UM feature because the BUM should not occur for $f_o < nf_{ce}$. In this case, the lower hybrid waves are weaker in the ion trajectories and hence the amplitude of the DM is down 5dB than the previous case. But the energy absorption of the pump field by the plasma in this case increases by 50% and the electron heating is higher, as can be seen in the field energy plots which are not displayed in this thesis[25].

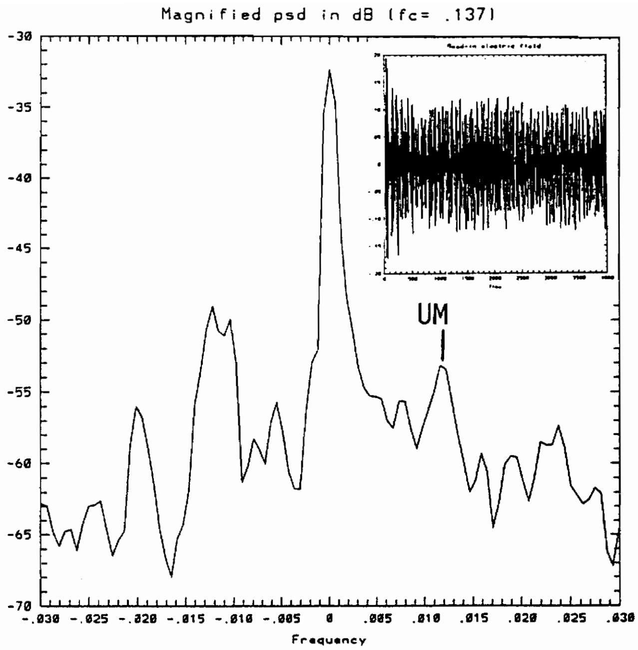
When the comprehensive spectrum in Figure 24(b) is compared with Figure 23(d), a number of differences are obvious. The first one is that a strong peak appears at a frequency about 0.07 in Figure 24(b). It is not close to neither the first nor the second electron cyclotron harmonics. Moreover, the power enhancement at higher harmonics reduces and the harmonic pattern becomes less clear. Also in this case, the frequency matching of the PDI is not as well defined as the reference case.

If the pump frequency decreases further to $\omega_o = 0.84$, the DM disappears from the spectrum, as shown in Figure 25(a). The settling time of the transient effect becomes longer in this case. It is evident from Figure 25(b) that there is no PDI occurring in the plasma. No lower hybrid waves are found in the ion trajectories. But, there exist some irregular power enhancements at certain frequencies above and below the pump, and the enhancement of higher harmonic power is less significant than the reference case.

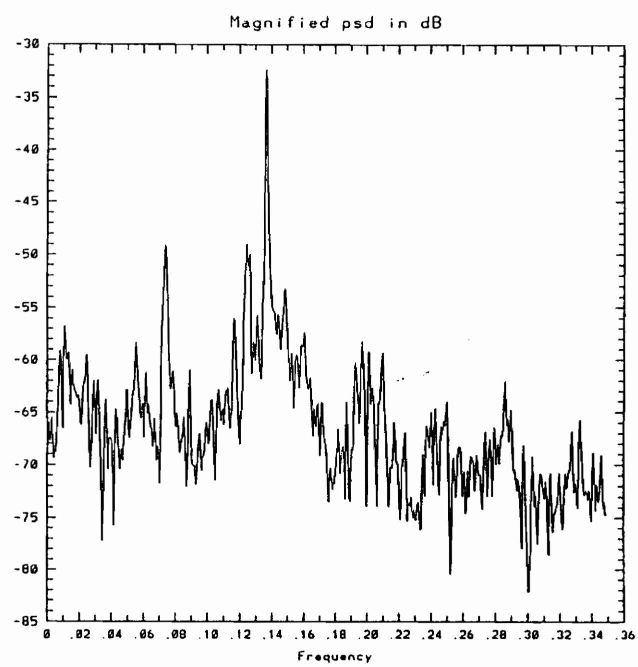
When the pump frequency increases from $\omega_o = 0.88$ to 0.89, the DM is still the dominant feature and its strength is comparable to the reference case, as shown in Figure 26(c). The low-frequency spike and the second DM in this case seem stronger (Figure 26(c) and (d)). From the trajectory

[24] The velocity is a derived quantity of the length and time. Its base reference can be in terms of either the base references of the length and time or the electron thermal velocity. All velocities shown in this chapter are understood that they are relative to v_{te}. However, the value of v_{te} may change from case to case. Comparison of the values of two velocities is only meaningful when they correspond to the same simulation setup.

[25] It is due to space limitation. A collection of the simulation results and output files is available from Dr. W.A. Scales.

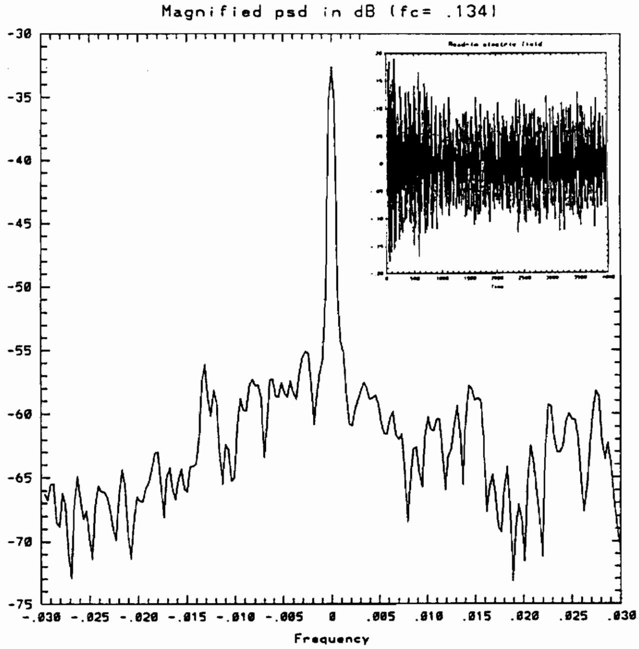


(a) Power spectrum for the time series from $t = 1000$ to $t = 2600$.

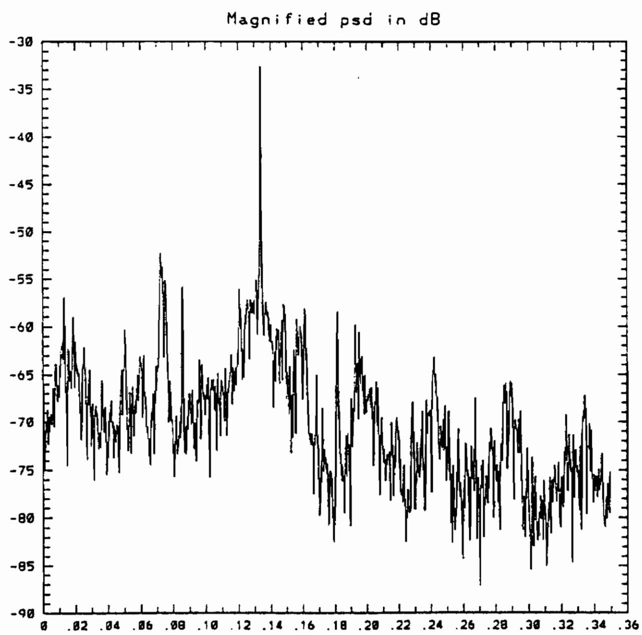


(b) Comprehensive power spectrum from zero frequency to 0.35. ($t = 1000$ to $t = 2600$)

Figure 24: Simulation results at $\omega_o = 0.86$, with *ecenter* = *ngpsd* = 136 and Region II parameters.



(a) Power spectrum for the time series from $t = 600$ to $t = 3800$.



(b) Comprehensive power spectrum from zero frequency to 0.35. ($t = 600$ to $t = 3800$)

Figure 25: Simulation results at $\omega_o = 0.84$, with $ecenter = ngpsd = 119$ and Region II parameters.

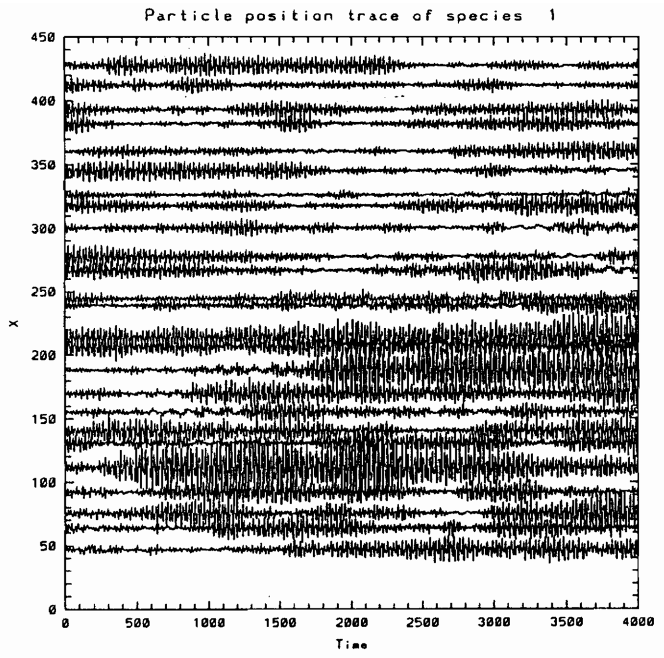
plots of electrons and ions (Figure 26(a) and (b)), the lower hybrid waves seem to propagate in both directions away from the pump region. This may be due to the excitation position in this case being further away from the left edge of the plasma. Note the reduced power enhancements at higher harmonics in Figure 26(d). The absorption of the pump energy is also less than the case $\omega_o = 0.88$.

From the fact that the lower hybrid waves propagate in both directions in Figure 26(b), we may want to know whether the generation of DM depends on the excitation position in the plasma. This will provide insight into how the finite size of the plasma affects the simulation results. Figure 27(a) to (c) show some of the simulation results for the pump frequency $\omega_o = 0.88$ with the pump region shifted to near the center of the plasma. In this case, the nominal plasma frequency ω_{peo} is changed to 0.9. The system length l is also changed to keep the charge and mass of the electrons same as the reference case. It is evident from Figure 27(a) and (b) that the lower hybrid waves propagate in both directions, but its propagation towards the higher density side is stronger and lasts for longer time. However, in Figure 27(c), the upper sideband peaking at a frequency of 0.075 above the pump is enhanced, although the DM is also enhanced. This upper sideband is most likely to be the UM feature because its frequency offset from the pump is less than that of the DM. Note that the higher order DM disappears in this case, and an additional peak shows up at a frequency of 0.022 above the pump.

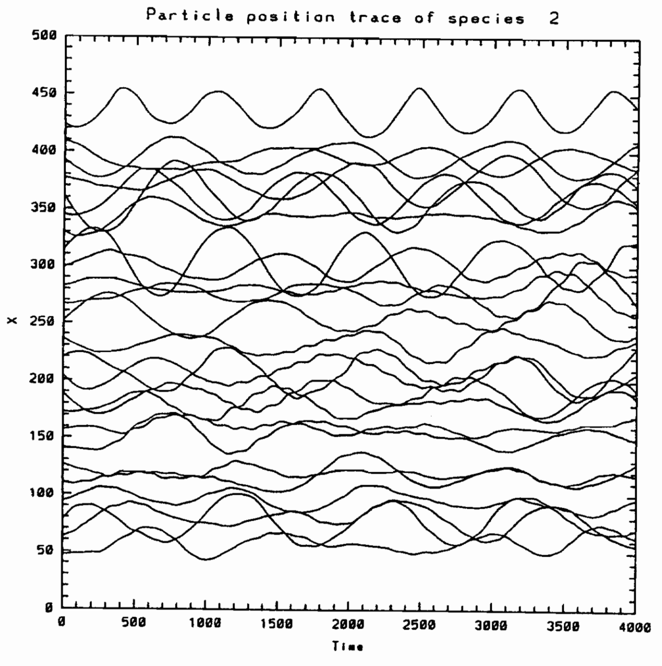
Fourth harmonic

Now, we go on to the cases slightly below $4f_{ce}$ and follow approximately the same order of the third harmonic cases. Figure 28(a) to (d) depict some important results at $\omega_o = 1.18$. Again, Region II parameters are used. The pump region in this case is at a position near the center of the plasma. Similar to the case in Figure 27, the lower hybrid waves propagate in both directions. However, the lower hybrid waves are considerably weaker in this case (Figure 28(b)), and the amplitude of the DM decreases (Figure 28(c)). But, in Figure 28(c), the UM is enhanced. This is the strongest UM ever seen in the SEE simulations. In Figure 28(d), it is evident that the PDI involving lower hybrid waves occurs in the plasma, and the energy of electron cyclotron harmonics above the upper hybrid frequency is enhanced. The absorption of the pump energy is nearly three times of the reference case. Also note that the DM peaks at a frequency of 0.009 below the pump, which is slightly less than that in the reference case. A very weak second DM appears in Figure 28(c).

Figure 29(a) shows the power spectrum of $\omega_o = 1.16$. The amplitudes of the DM and UM are nearly equal. The settling time of the transient effects is longer than the case $\omega_o = 1.18$. In Figure 29(b), a signature of PDI can be seen. Similar to the case $\omega_o = 0.86$, a significant power enhancement occurs at a frequency around 0.14. It is close to the third electron cyclotron harmonic. The enhancement of higher harmonic power is comparable to the previous case. When the pump

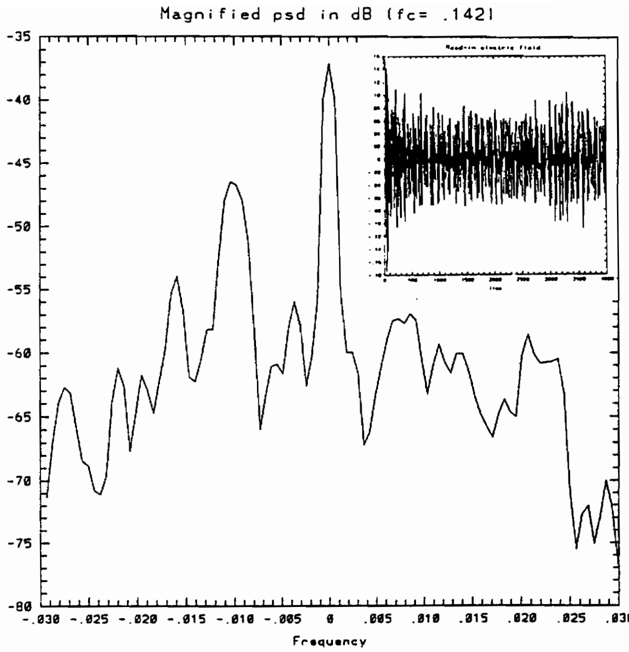


(a) Electron trajectories.

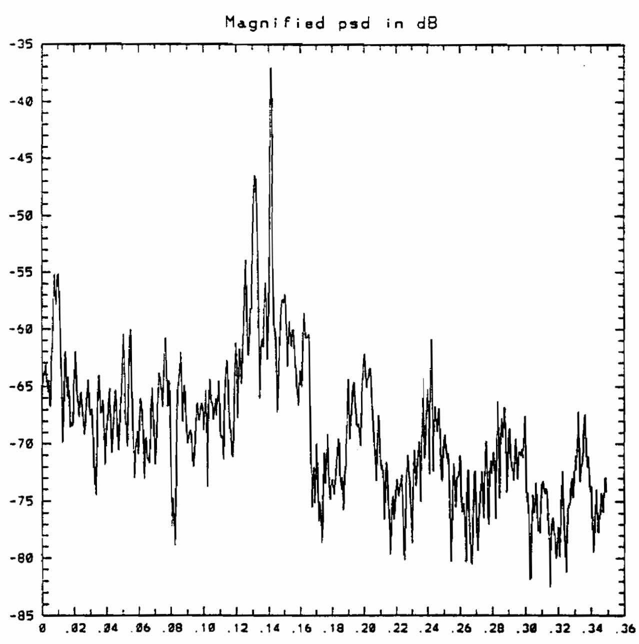


(b) Ion trajectories.

Figure 26: Simulation results at $\omega_o = 0.89$, with $ecenter = ngpsd = 163$ and Region II parameters.

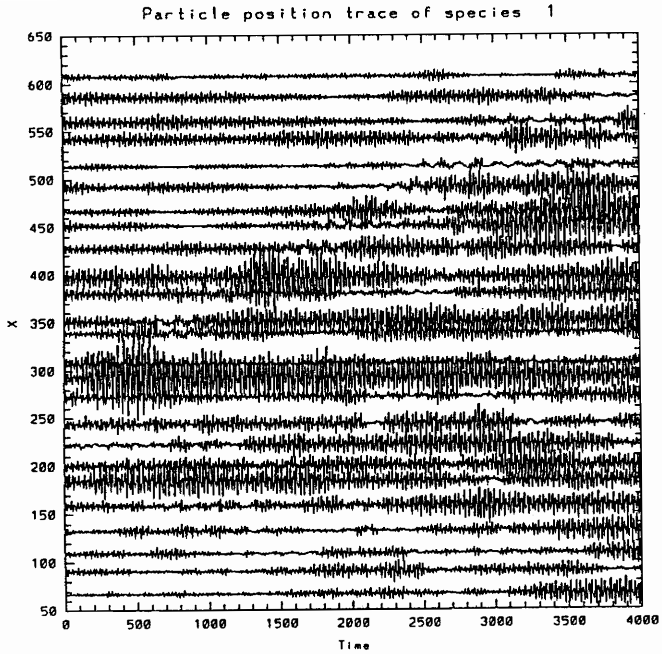


(c) Power spectrum for the time series from $t = 2000$ to $t = 3600$.

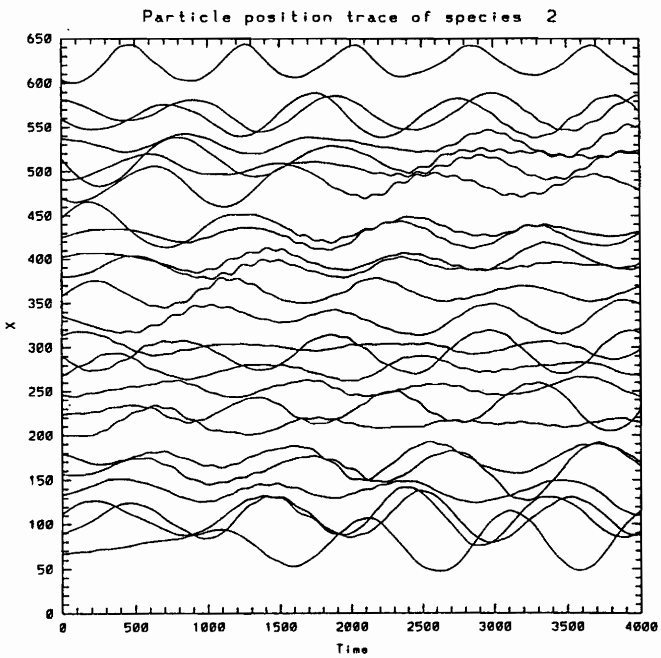


(d) Comprehensive power spectrum from zero frequency to 0.35. ($t = 2000$ to $t = 3600$)

Figure 26 (continued).

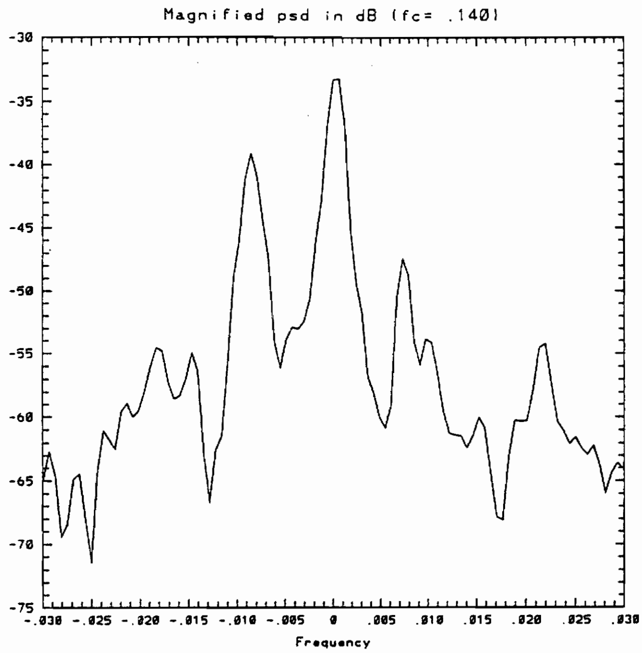


(a) Electron trajectories.



(b) Ion trajectories.

Figure 27: Simulation results at $\omega_o = 0.88$, with $lll = 106.8$, $\omega_{peo} = 0.9$, $ecenter = ngpsd = 393$ and $dfn = 1.314$.



(c) Power spectrum for the time series from $t = 1000$ to $t = 2600$.

Figure 27 (continued).

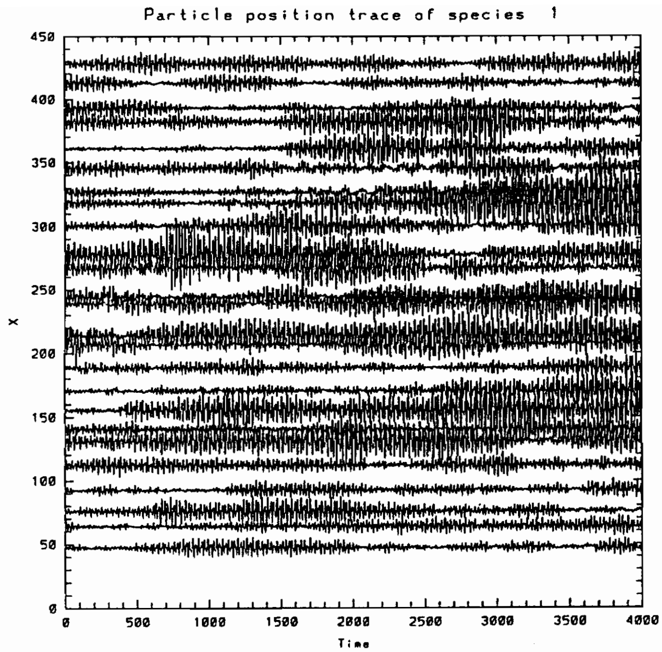
frequency decreases to $\omega_o = 1.14$, both the DM and UM almost disappear (Figure 30) because the PDI is very weak in this case. But the energy absorption inside the plasma are higher than the previous two cases.

When the pump frequency steps close to $4f_{ce}$, the pump is damped. Figure 31(a) shows the power spectrum of $\omega_o = 1.19$. Both the DM and UM are present. The strength of the DM and UM are comparable to the case $\omega_o = 1.18$. Note that the offset frequency of the DM from the pump decreases and the offset frequency of the UM increases in this case. Moreover, the bandwidth of the DM shrinks to 0.003. Again, we do not know what causes the regular structure centered at a frequency of 0.02 above the pump. As expected, the PDI is weak in this case (Figure 31(b)). The enhancement at higher cyclotron harmonics is comparable to the case $\omega_o = 1.18$.

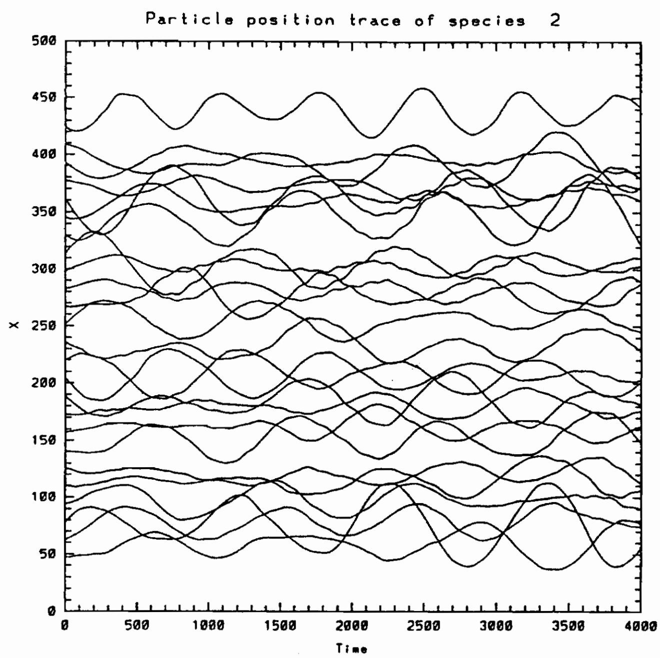
Figure 32 shows the case when the pump frequency is very close to $4f_{ce}$ ($\omega_o = 1.195$). The UM feature almost disappears. The PDI is weaker than the previous case.

Different pump positions

To investigate the effect of pump position in the plasma, two simulations were performed at $\omega_o = 1.18$. Figure 33(a) corresponds to the case which the pump position is close to the left edge of

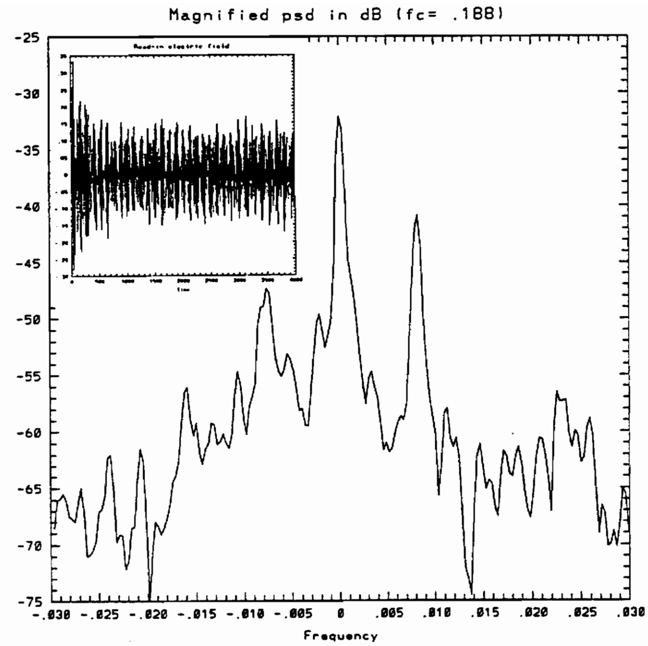


(a) Electron trajectories.

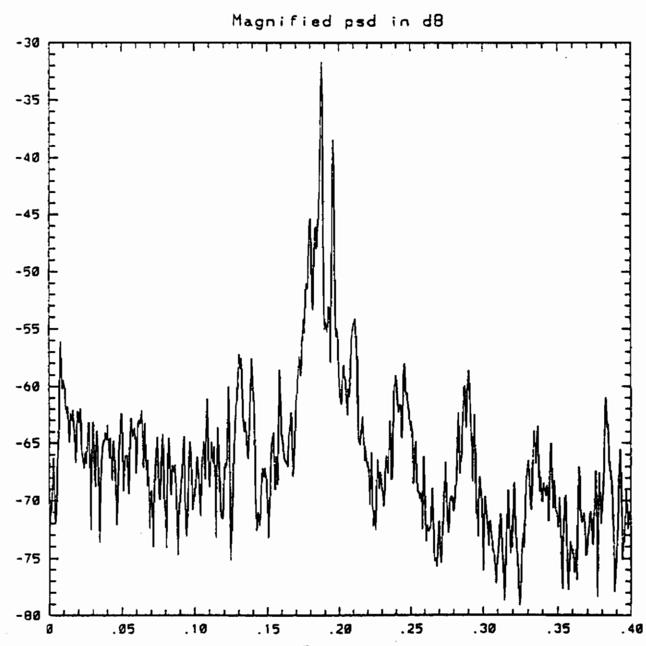


(b) Ion trajectories.

Figure 28: Simulation results at $\omega_o = 1.18$, with $ecenter = ngpsd = 463$ and Region II parameters.

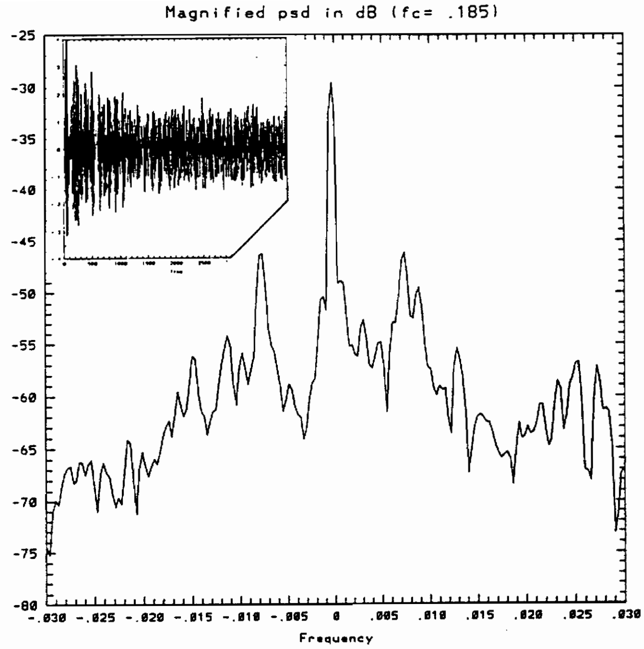


(c) Power spectrum for the time series from $t = 600$ to $t = 3800$.

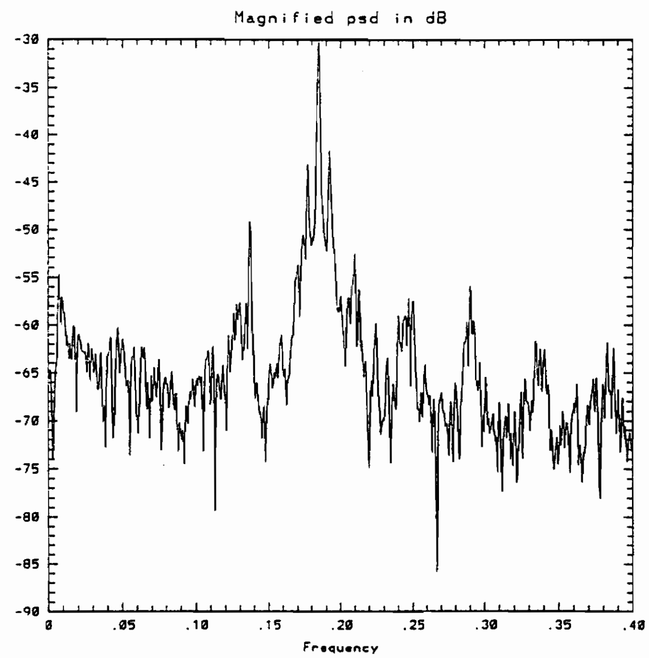


(d) Comprehensive power spectrum from zero frequency to 0.4. ($t = 400$ to $t = 2000$)

Figure 28 (continued).

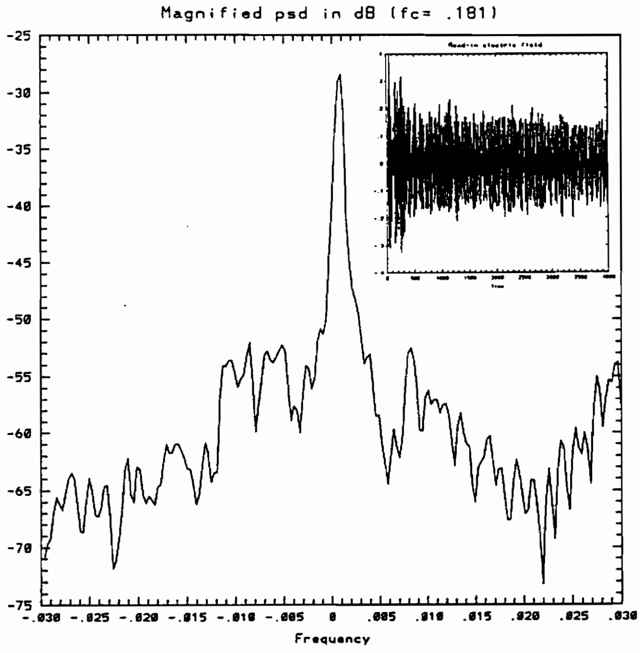


(a) Power spectrum for the time series from $t = 600$ to $t = 3800$.



(b) Comprehensive power spectrum from zero frequency to 0.4. ($t = 600$ to $t = 2200$)

Figure 29: Simulation results at $\omega_o = 1.16$, with $ecenter = ngpsd = 440$ and Region II parameters.



Power spectrum for the time series from $t = 600$ to $t = 3800$.

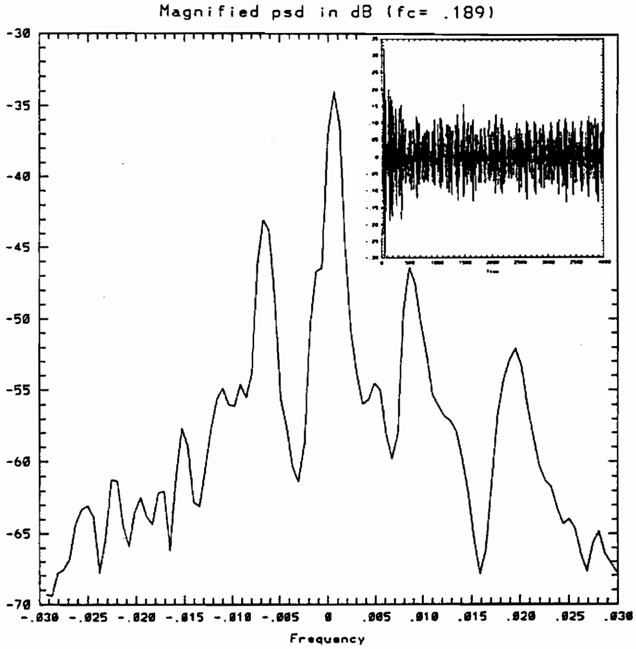
Figure 30: Simulation results at $\omega_o = 1.14$, with $ecenter = ngpsd = 417$ and Region II parameters.

the plasma. The resulting spectrum is very different from Figure 28(c). The UM is suppressed in this case, and the DM is weaker and is split into two halves. When the excitation position moves near the right edge of the plasma (Figure 33(b)), the UM is strong and the DM is split into many higher order DMs.

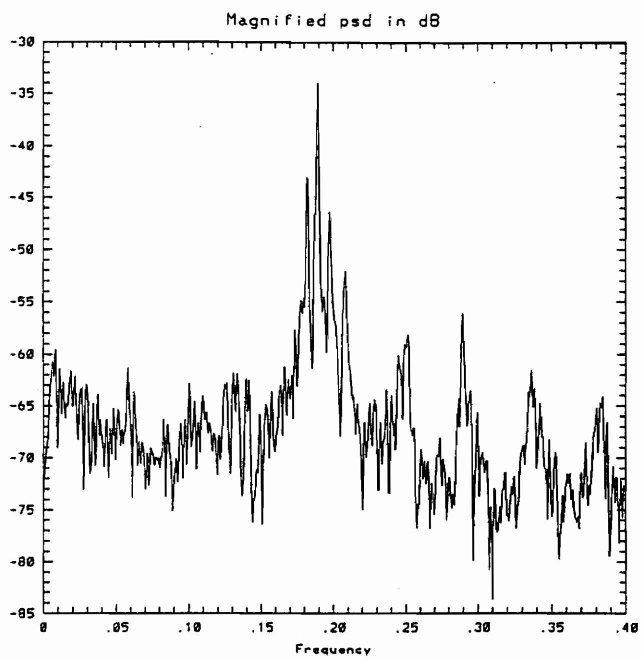
Other harmonics

Besides the third and fourth harmonics, the DM and UM features are present in cases where the pump frequencies are slightly below the second and fifth electron cyclotron harmonics. For the second harmonic, Region I parameters are used. Figure 34(a) shows a very clear development of the DM, with a very weak UM around a frequency offset of 0.005 above the pump. A very weak second DM also appears in Figure 34(a). In this case, the pump region is near the left edge of the plasma. As seen from Figure 34(b), the PDI is strongly developed, and the energy at higher harmonics is slightly enhanced.

Figure 35(a) to (d) depict some important simulation results for $\omega_o = 1.48$ (the fifth harmonic). Region III parameters are used. The electron heating by the propagation of lower hybrid waves is clear in Figure 35(a). The pump region is approximately 140 grid cells below the plasma center and

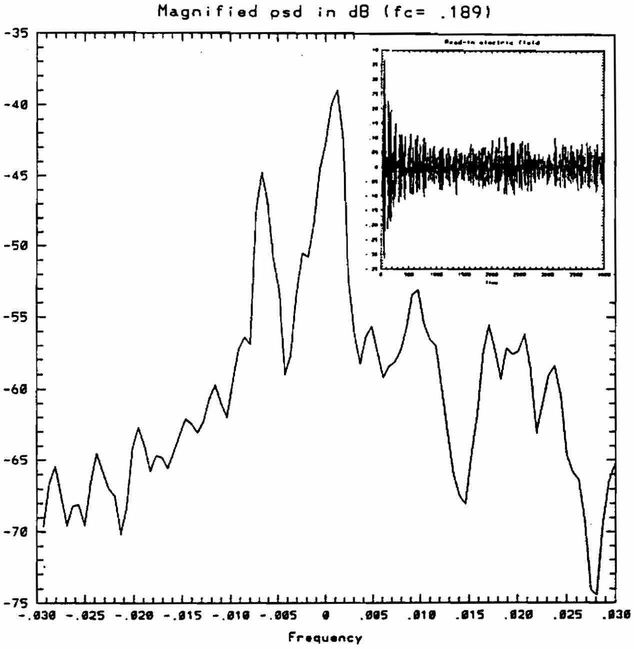


(a) Power spectrum for the time series from $t = 500$ to $t = 2100$.



(b) Comprehensive power spectrum from zero frequency to 0.4. ($t = 500$ to $t = 2100$)

Figure 31: Simulation results at $\omega_o = 1.19$, with $ecenter = ngpsd = 475$ and Region II parameters.



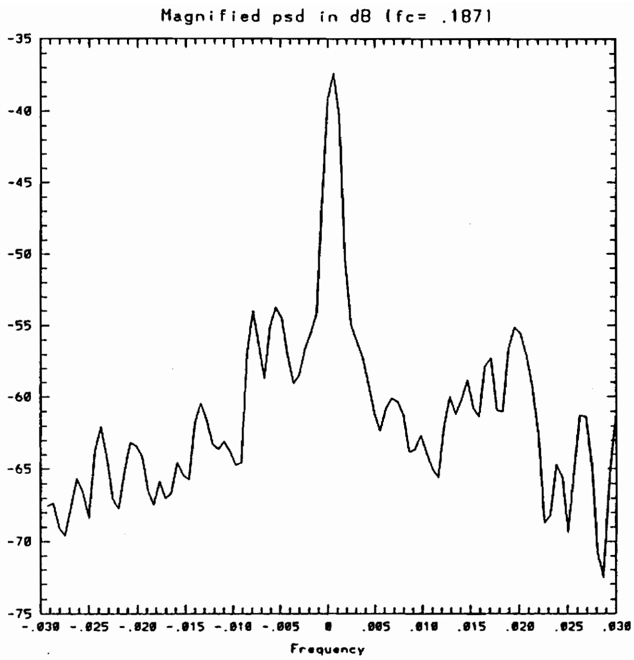
Power spectrum for the time series from $t = 1000$ to $t = 2600$.

Figure 32: Simulation results at $\omega_o = 1.195$, with *ecenter* = *ngpsd* = 481 and Region II parameters.

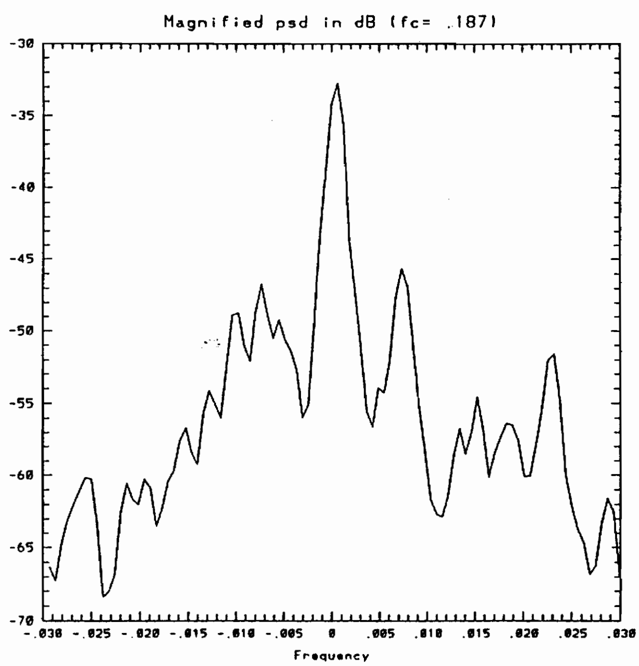
the pump field is effective over 121 grid cells in this case. The DM and UM are both present in the power spectrum (Figure 35(c)). Note that the frequency offset of the UM from the pump is larger than the DM. In Figure 35(d), the PDI is quite strong, and the power enhancement at the eighth cyclotron harmonic is significant.

Effect of mass ratio

In the following, the dependence of the SEE spectrum on the mass ratio, observation point, and electron and ion temperatures is investigated. Figure 36 and 37 depict two power spectra which are simulated with the same parameters as the reference case, except $K = 100$ and $K = 225$ respectively. The DM feature clearly shows its dependence on the ion mass. It peaks at a frequency of 0.006 below the pump in Figure 36 and at 0.004 below the pump in Figure 37. The lower hybrid frequencies for $K = 100$ and $K = 225$ are 0.0045 and 0.003 respectively. The frequency offset of the DM from the pump is approximately proportional to $1/\sqrt{K}$ for $K = 49, 100, 225$. Note that the frequency resolution in Figure 37 cannot accurately show the high-frequency cutoff of the DM exactly at the lower hybrid frequency. Also, only three and a half of ion gyromotion cycles are included in the simulation with $K = 100$ and even fewer in the simulation with $K = 225$. This may cause the

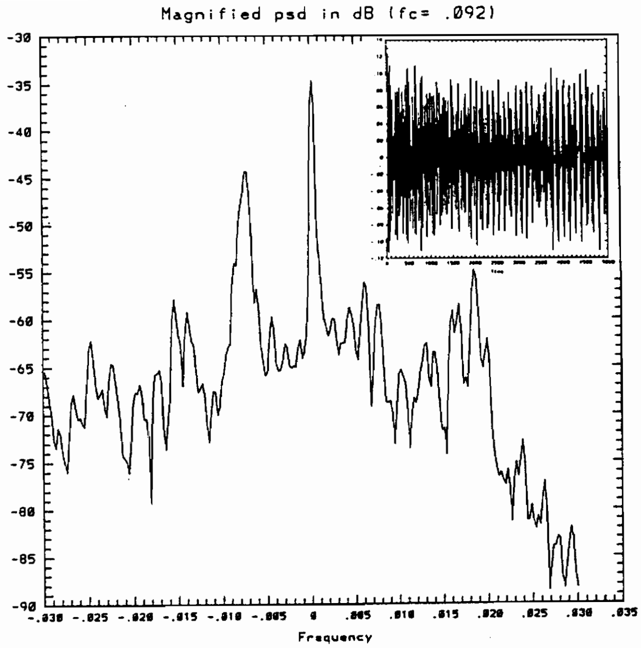


(a) Pump region close to the left edge of the plasma. ($lll = 33.8$, $\omega_{peo} = 1.6$, $ecenter = ngpsd = 161$ and $dfn = 1.314$)

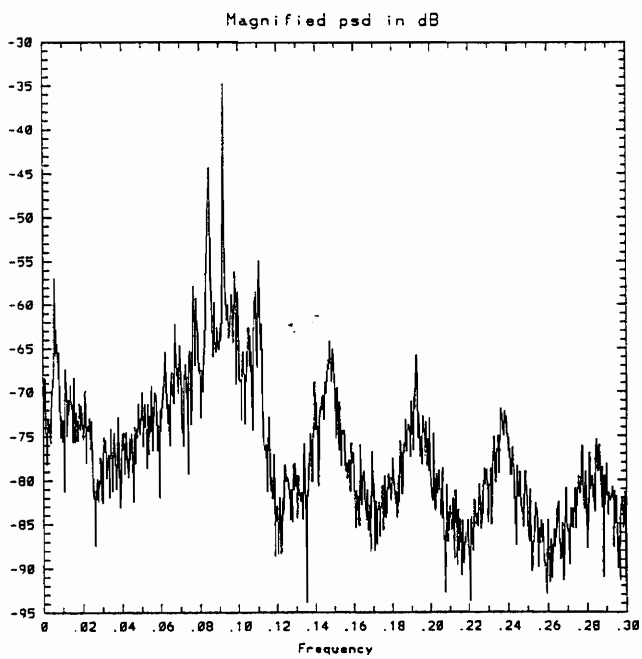


(b) Pump region close to the right edge of the plasma. ($lll = 100$, $\omega_{peo} = 1.0$, $ecenter = ngpsd = 707$ and $dfn = 1.314$)

Figure 33: Simulation results at $\omega_o = 1.18$ for two different excitation positions. Both spectra are taken from $t = 1000$ to $t = 2600$.

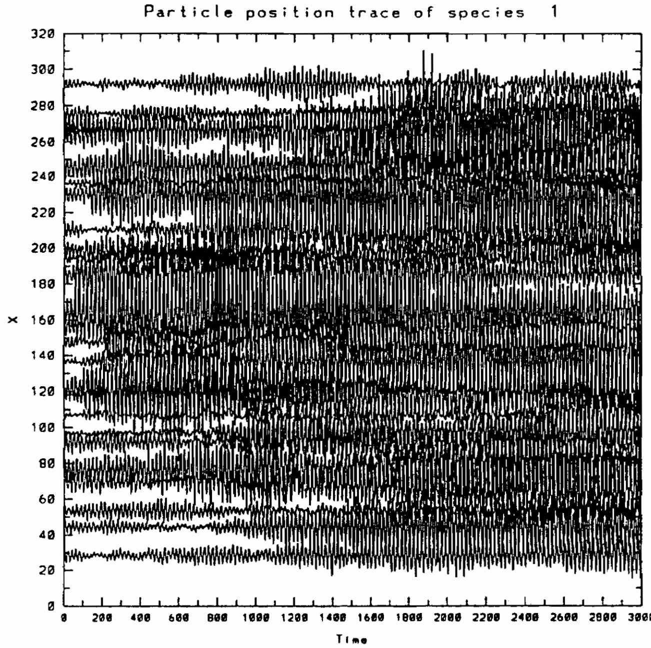


(a) Power spectrum for the time series from $t = 750$ to $t = 4750$.

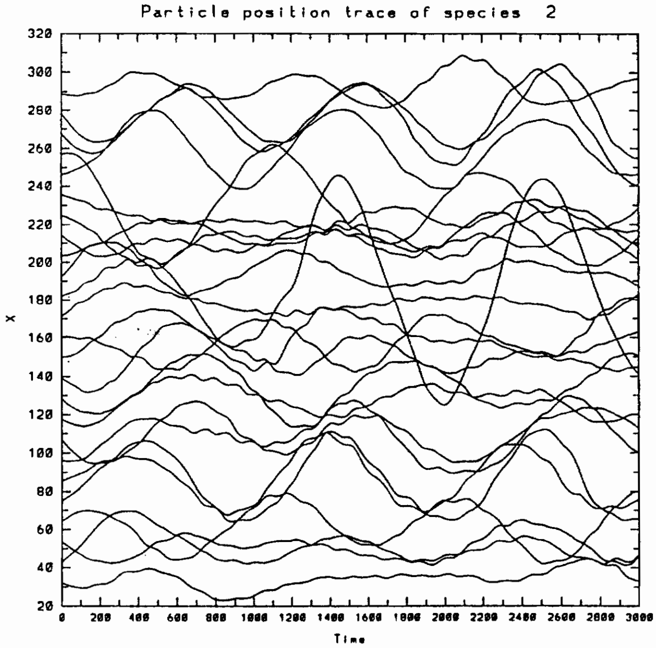


(b) Comprehensive power spectrum from zero frequency to 0.3. ($t = 750$ to $t = 4750$)

Figure 34: Simulation results at $\omega_o = 0.58$, with $ecenter = ngpsd = 107$ and Region I parameters.

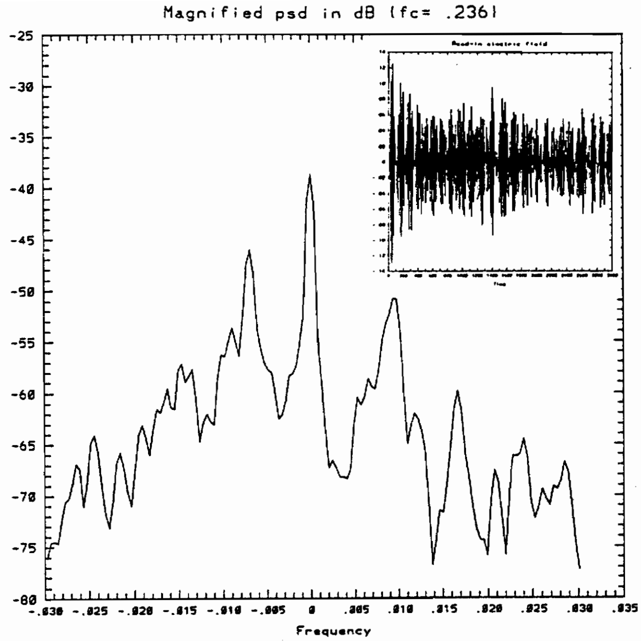


(a) Electron trajectories.

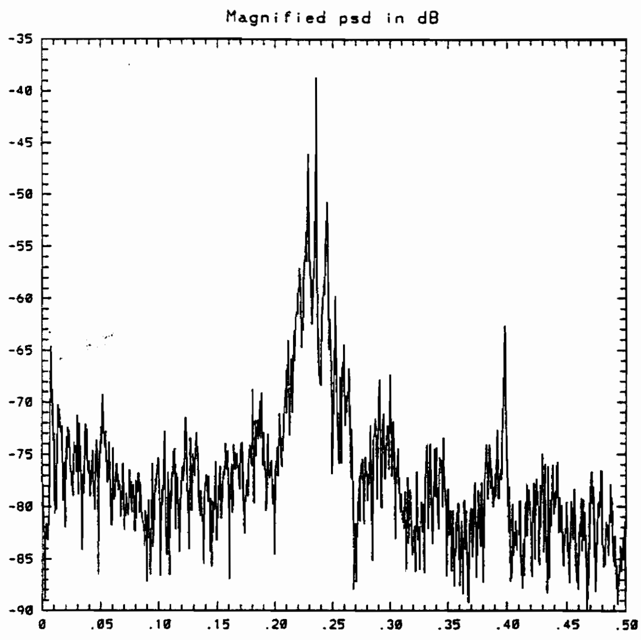


(b) Ion trajectories.

Figure 35: Simulation results at $\omega_o = 1.48$, with $ecenter = ngpsd = 373$ and Region III parameters.



(c) Power spectrum for the time series from $t = 450$ to $t = 2850$.



(d) Comprehensive power spectrum from zero frequency to 0.5. ($t = 450$ to $t = 2850$)

Figure 35 (continued).

strange spectral shape in Figure 36.

Different observation positions

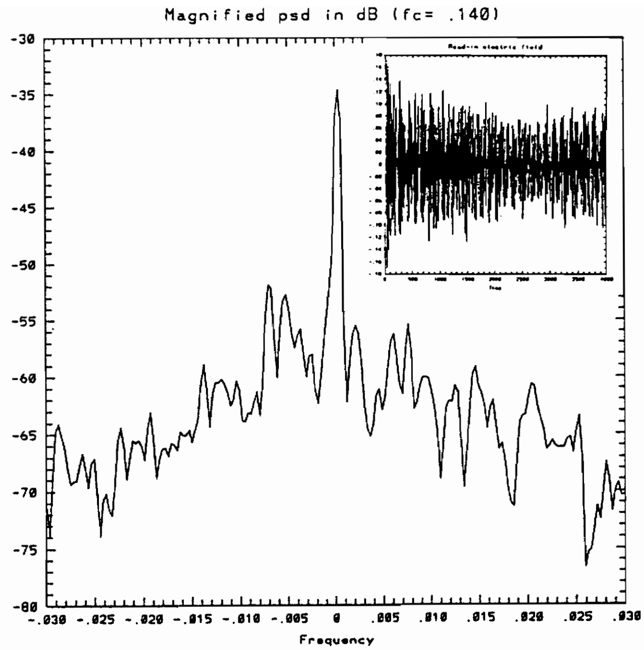
Since the lower hybrid wave propagates away from the pump region, we may be interested to know whether the same phenomenon occurs with the other decayed wave, i.e. the one carries the DM frequency. Figure 38(a) to (d) show the power spectra at four different observation locations. At $ng = 104$ (50 grid cells away from *ecenter*) the pump amplitude is reduced and the DM spreads out (Figure 38(a)). In Figure 38(b), the DM almost disappears. Note that $ng = 124$ is just at the left border of the pump region. Figure 38(c) shows the power spectrum at the upper hybrid point of the DM, i.e. the upper hybrid frequency equals the frequency of the DM peak at $ng = 146$. It shows no particular enhancement of the DM feature. Figure 38(d) depicts the power spectrum on the right side of the pump region. The DM amplitude is slightly decreased and its bandwidth is increased. An interesting case of the effect of observation point is shown in Figure 39. The amplitude of the DM well exceeds the pump at 100 grid cells below the pump region. This is a very good evidence to prove that the decay wave carrying the DM frequency can propagate far away from the pump region. From the above observations, it raises two questions which have never considered in the existing SEE theories.

1. What causes the variations of the DM amplitude along the plasma? Is the structure periodic and related to some type of boundary conditions?
2. Where does the backward mode conversion of the DM occur inside the plasma? It is quite sure that it should not happen in the regions above the pump region because the resulting electromagnetic wave cannot propagate through the pump region.

Effect of electron and ion temperatures

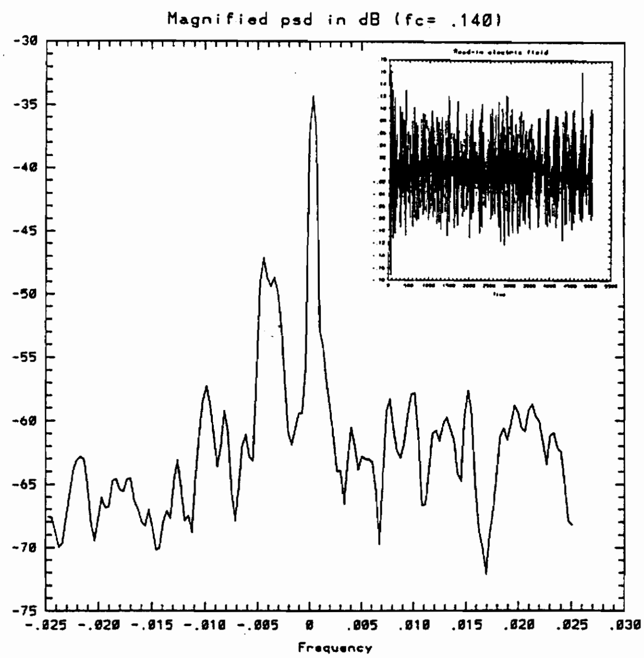
The effect of electron and ion temperatures on the power spectrum is investigated in the following. Figure 40(a) to (c) correspond to the simulation with a reduced electron temperature. As expected, the electron trajectories are much quieter (Figure 40(a)). It clearly demonstrates the electron heating triggered by an outgoing wave. The ripples produced by lower hybrid waves in Figure 40(b) seem to be weaker than the reference case. Nevertheless, the DM is well developed in Figure 40(c). The offset frequency of the DM from the pump is roughly the same as the reference case.

Figure 41 shows the case with doubled electron thermal velocity. The offset frequency of the DM from the pump increases by about 0.001 compared with the reference case. Figure 42 depicts the case with halved ion thermal velocity. The DM offset frequency decreases by 0.0015 in this case. Also, the amplitude of the DM is slightly enhanced. If the ion thermal velocity is doubled ($v_{ti} = 0.143$),



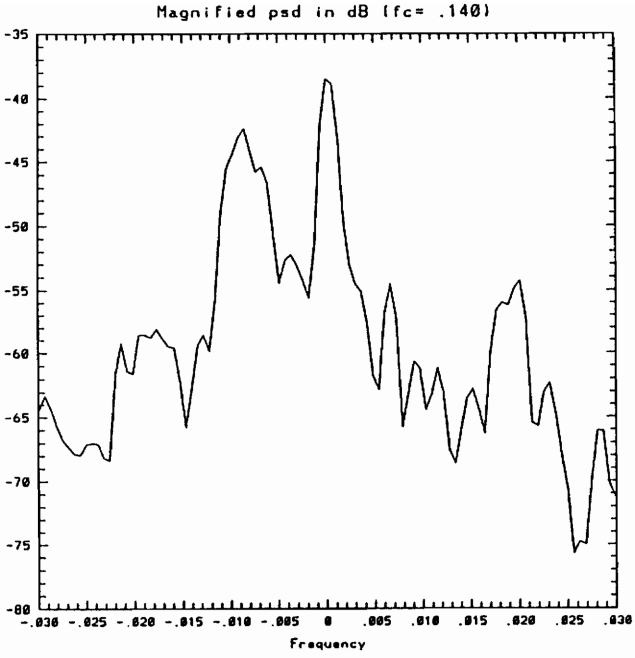
Power spectrum for the time series from $t = 600$ to $t = 3800$.

Figure 36: Simulation result at $\omega_o = 0.88$, with $ecenter = ngpsd = 157$, $K = 100$ and Region II parameters.

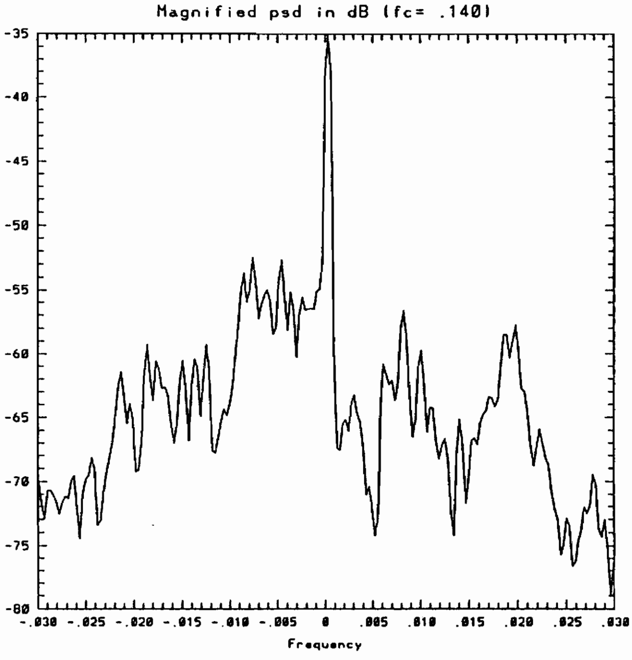


Power spectrum for the time series from $t = 600$ to $t = 3800$.

Figure 37: Effect of mass ratio on power spectrum ($\omega_o = 0.88$, $ecenter = ngpsd = 159$, $K = 225$ and Region II parameters).

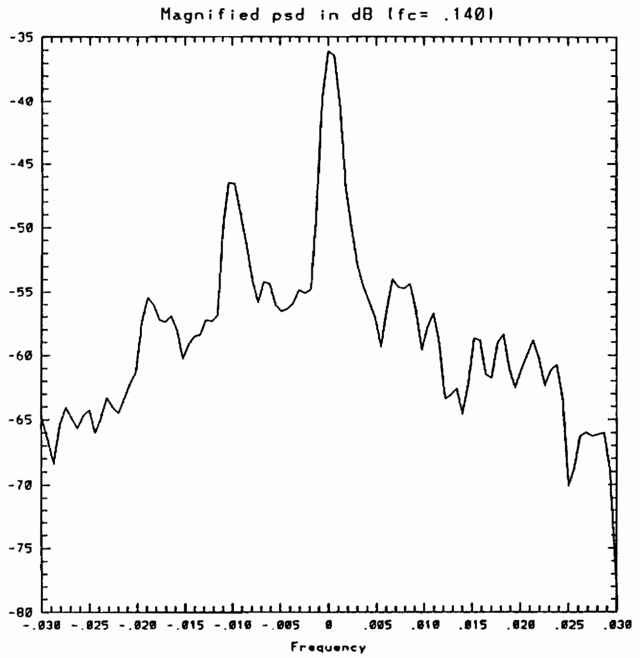


(a) Power spectrum at $ng = 104$. ($t = 1000$ to $t = 2600$)

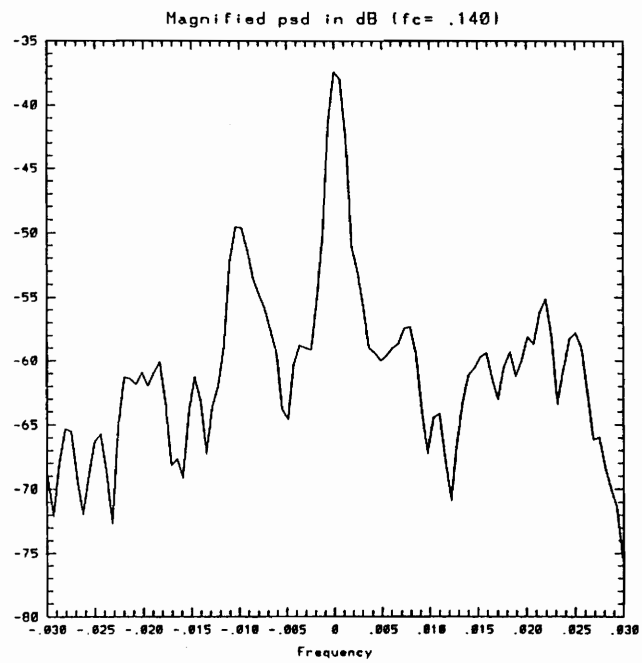


(b) Power spectrum at $ng = 124$. ($t = 600$ to $t = 3800$)

Figure 38: Effect of observation point on power spectrum ($\omega_o = 0.88$, $ecenter = 154$ and Region II parameters).

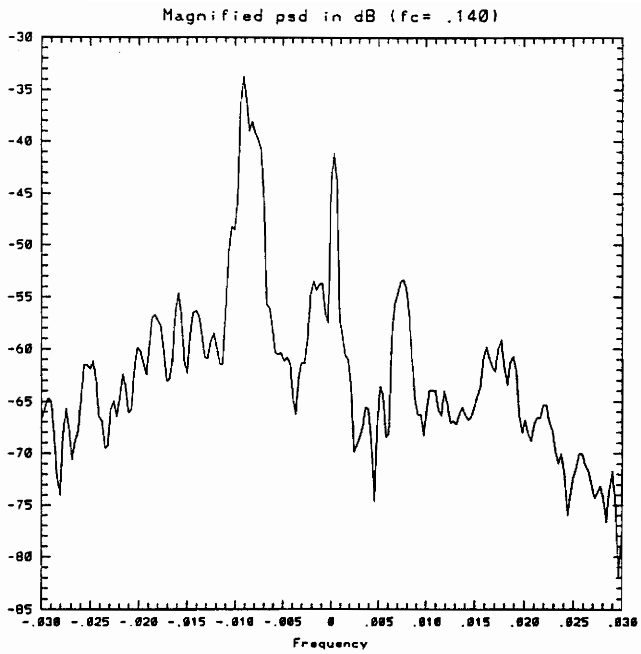


(c) Power spectrum at $ng = 146$. ($t = 1000$ to $t = 2600$)



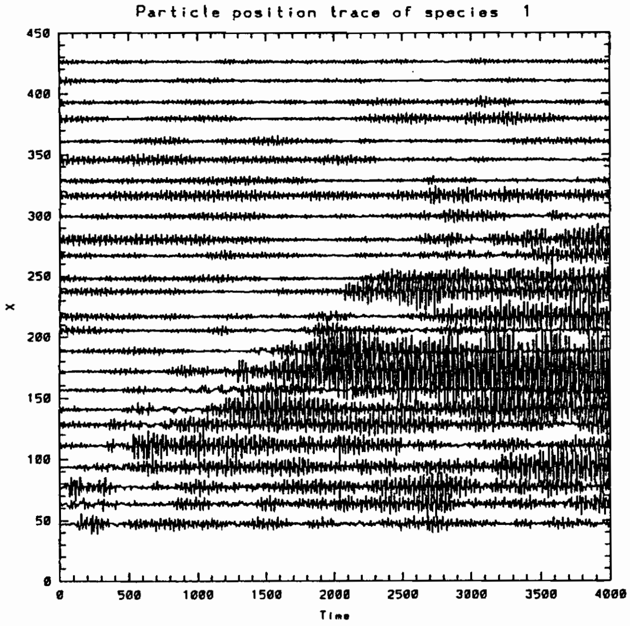
(d) Power spectrum at $ng = 204$. ($t = 1000$ to $t = 2600$)

Figure 38 (continued).

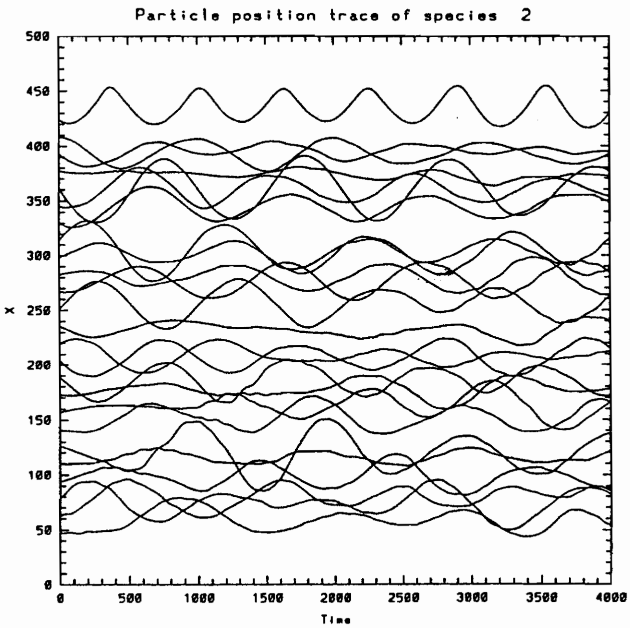


Power spectrum at $ng = 293$. ($t = 600$ to $t = 3800$)

Figure 39: Effect of observation point on power spectrum ($\omega_o = 0.88$, $ecenter = 393$ and parameters similar to the case in Figure 27).

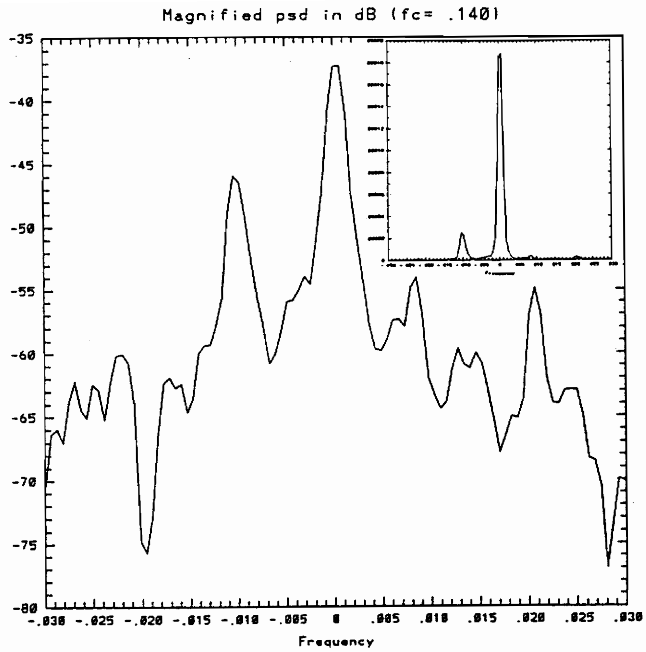


(a) Electron trajectories.



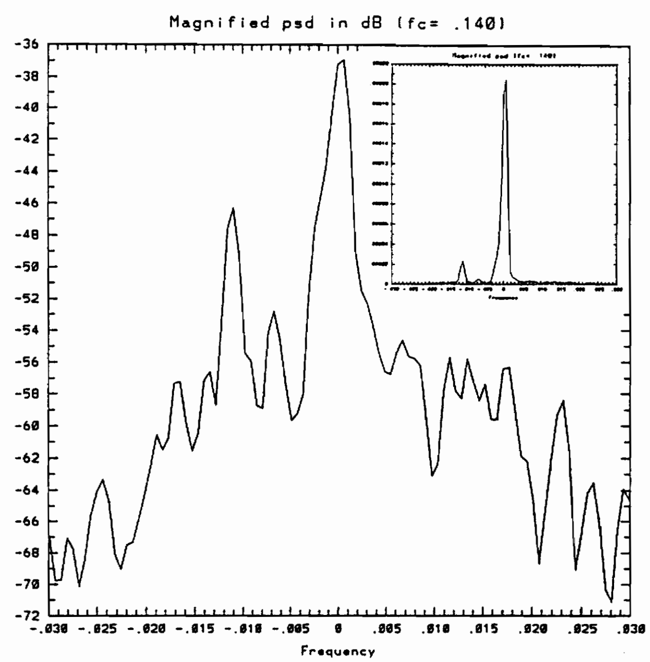
(b) Ion trajectories.

Figure 40: Effect of electron temperature on power spectrum ($\omega_o = 0.88$, $ecenter = 154$, $v_{te} = 0.5$ and Region II parameters).



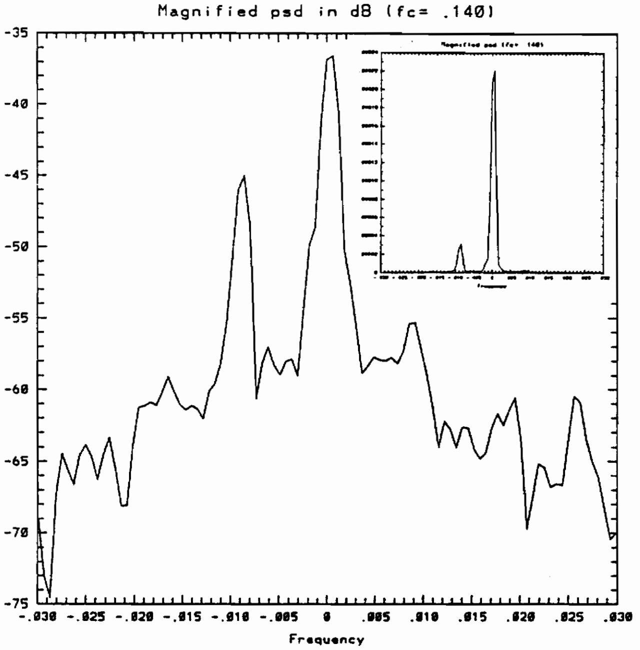
(c) Power spectrum ($t = 1000$ to $t = 2600$). The power spectrum at the right corner is plotted in linear scale.

Figure 40 (continued).



Power spectrum ($t = 1000$ to $t = 2600$).

Figure 41: Effect of electron temperature on power spectrum ($\omega_o = 0.88$, $ecenter = 154$, $v_{te} = 2$ and Region II parameters).



Power spectrum ($t = 1000$ to $t = 2600$).

Figure 42: Effect of ion temperature on power spectrum ($\omega_o = 0.88$, $ecenter = 154$, $v_{ti} = 0.0357$ and Region II parameters). Note that the ion temperature used in the reference is 0.0714.

the DM feature disappears because the PDI threshold cannot be exceeded by the pump. The PDI threshold, however, shows no significant dependence on the electron temperature.

5.3.2 Pump frequencies equal to nf_{ce}

We will show two cases which the pump frequencies are exactly at $3f_{ce}$ and $4f_{ce}$. The pump wave in general is heavily damped in both cases, and the absorption of the pump energy is very low. For $\omega_o = 0.9$, the electrons are heated only in the narrow region around the pump position (Figure 43(a)) and no propagation of lower hybrid waves is observed in the ion trajectories (Figure 43(b)). The pump amplitude is down about 10dB compared with the reference case. As expected, the DM and UM features are quenched in Figure 43(c) because of the suppression of PDI. The energy at higher cyclotron harmonics is only slightly enhanced ($\sim 5dB$), as shown in Figure 43(d).

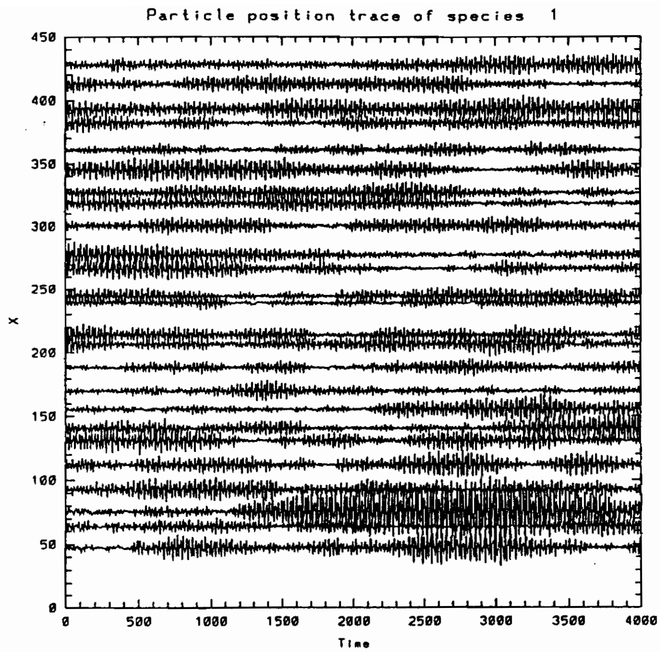
For the fourth electron cyclotron harmonic, less electron heating and pump energy absorption are resulted (Figure 44(a)). The ion trajectories stay quiet as in the third harmonic case (Figure 44(b)). The pump amplitude is slightly less than the third harmonic case. However, the DM and UM show up, although weaker, in the spectrum (Figure 44(c)). The pump amplitude is more heavily damped than the previous case. From Figure 44(d), the PDI seems to have occurred in the plasma, and additional energy enhancement can be observed at a frequency around 0.14 and at higher harmonics. This may suggest that the PDI threhold somewhat depends on the harmonic number.

5.3.3 Pump frequencies slightly above nf_{ce}

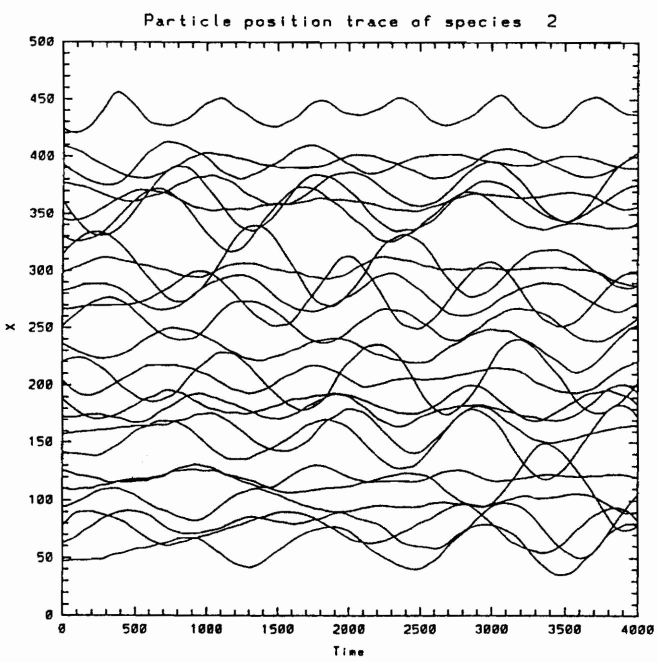
Third harmonic

Experiments show that the BUM feature appears in the SEE power spectrum when the pump frequency is slightly higher than nf_{ce}. Figure 45(a) and (b) depict the electron and ion trajectories for the case $\omega_o = 0.91$ in the simulation. They are very different from the reference case where $\omega_o = 0.88$. No propagating ripples of lower hybrid waves are present in the ion trajectories, and the electron heating is localized to the pump region and below. The settling time of transient effects is similar to the reference case (Figure 45(c)). The relatively broad upper sideband in Figure 45(c) is suspected to be the BUM[26]. Its bandwidth is about 0.009 and it has no well-defined peak. The unknown feature also appears at a frequency of 0.017 above the pump. The power enhancement at other frequencies is different from the reference case. In Figure 45(d), one cannot find a signature of PDI at low frequency. Besides the power enhancements at higher cyclotron harmonics (more than the reference case), the power levels at slightly above the zeroth and the first electron cyclotron

[26] We will use the name "BUM" for similar spectra in the following simulation results.

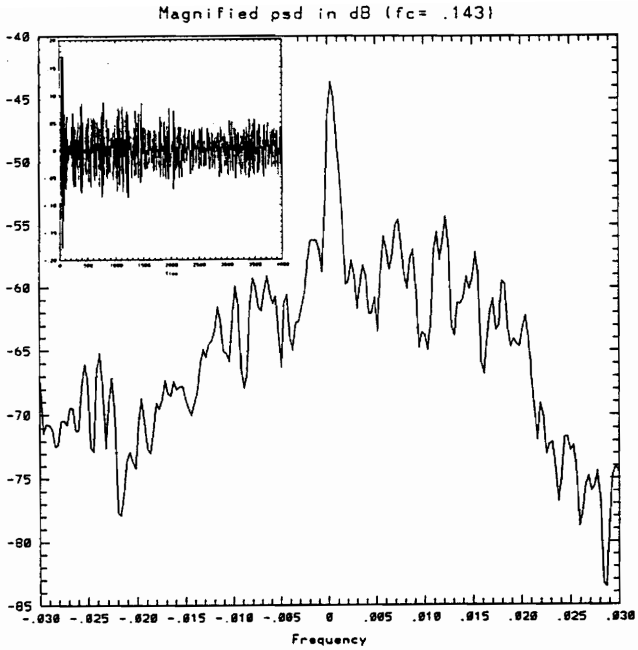


(a) Electron trajectories.

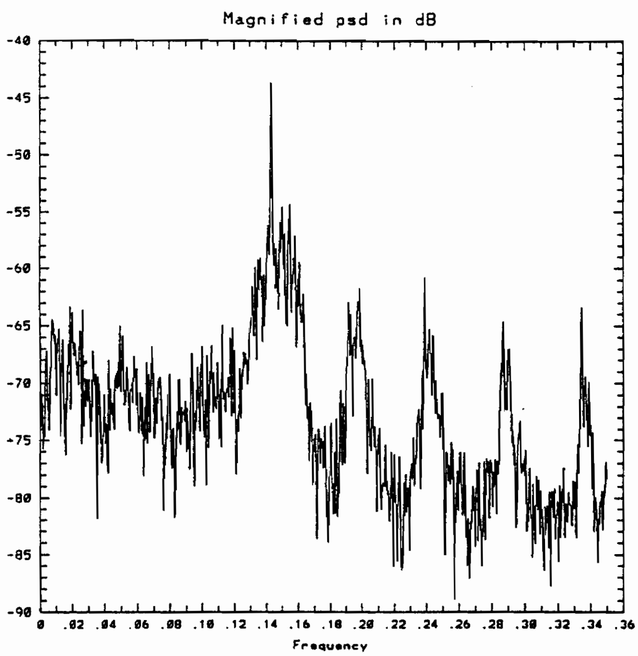


(b) Ion trajectories.

Figure 43: Simulation results at $\omega_o = 0.9$, with $ecenter = ngpsd = 172$ and Region II parameters.

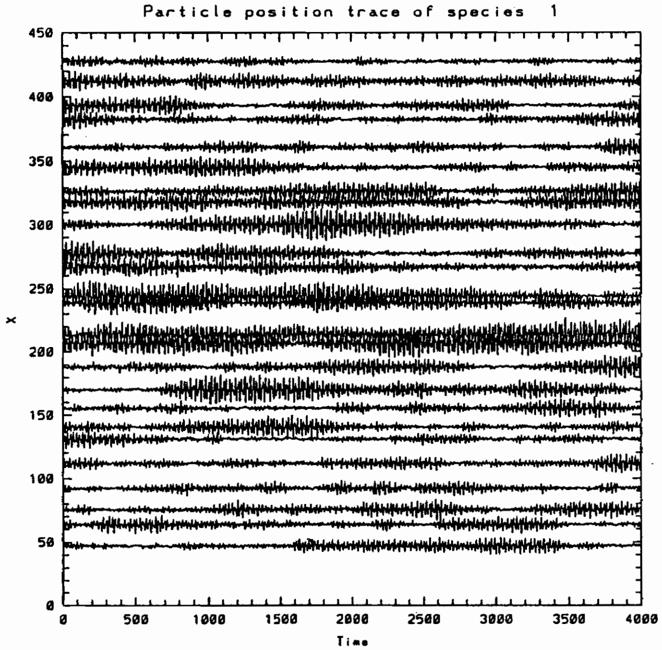


(c) Power spectrum for the time series from $t = 600$ to $t = 3800$.

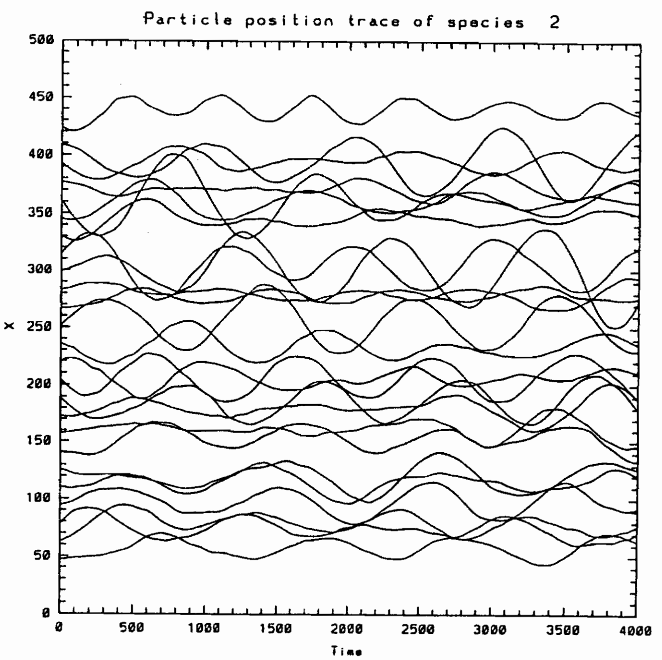


(d) Comprehensive power spectrum from zero frequency to 0.35. ($t = 600$ to $t = 3800$)

Figure 43 (continued).

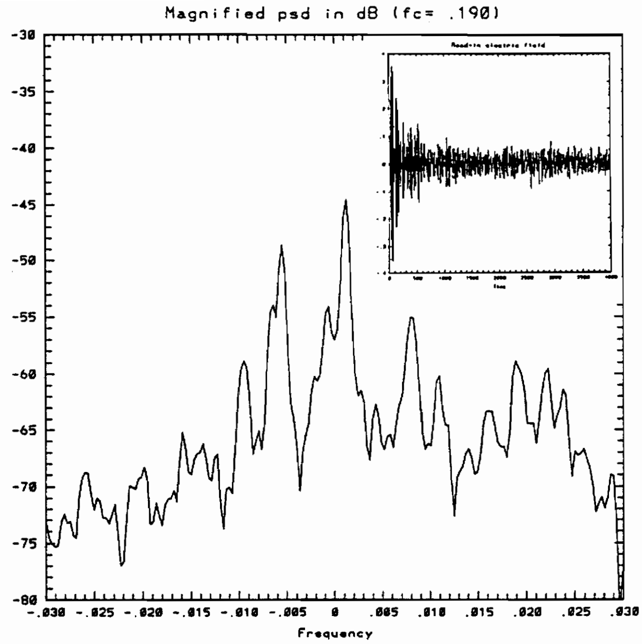


(a) Electron trajectories.

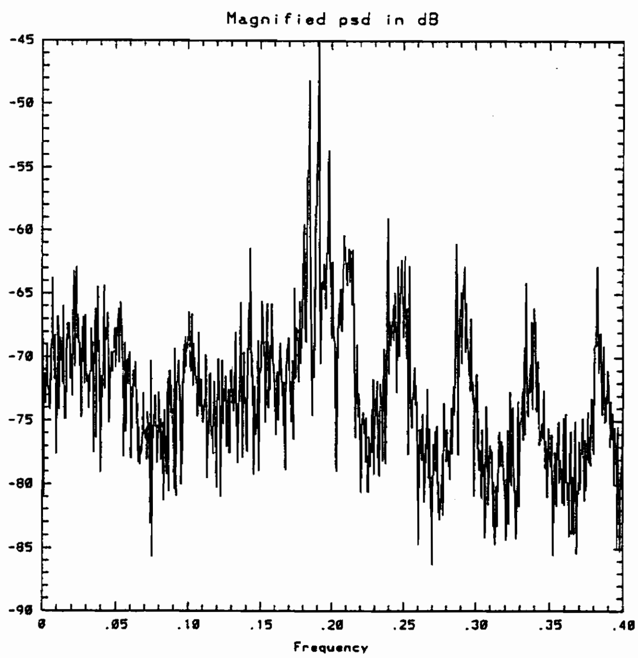


(b) Ion trajectories.

Figure 44: Simulation results at $\omega_o = 1.2$, with *ecenter* = *ngpsd* = 487 and Region II parameters.

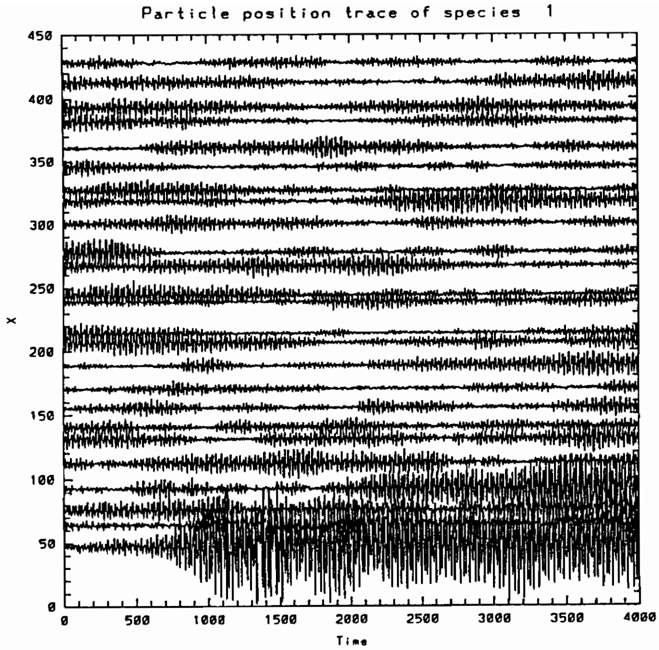


(c) Power spectrum for the time series from $t = 600$ to $t = 3800$.

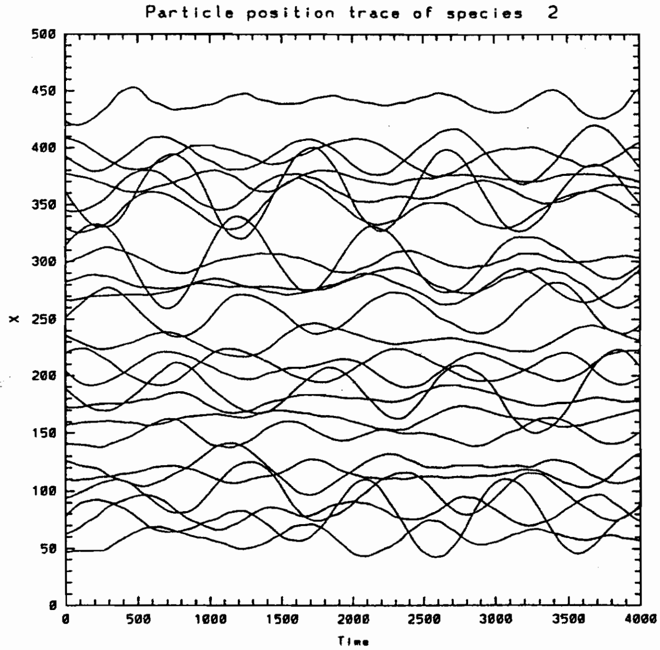


(d) Comprehensive power spectrum from zero frequency to 0.4. ($t = 800$ to $t = 4000$)

Figure 44 (continued).

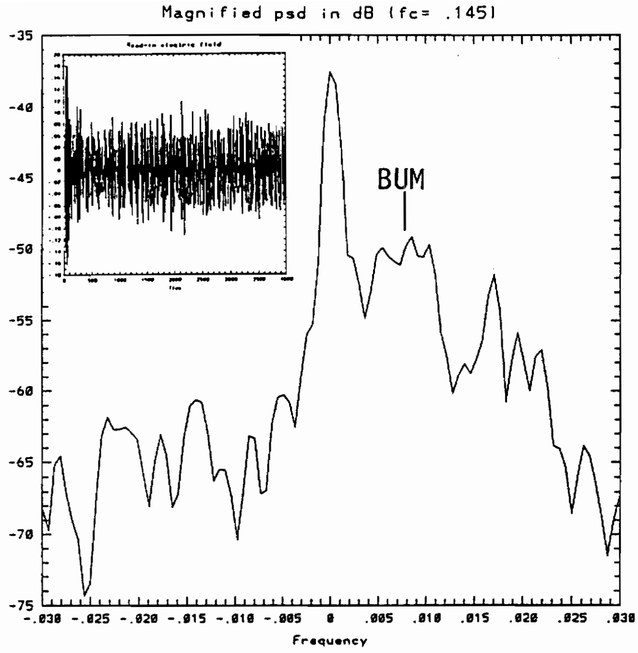


(a) Electron trajectories.

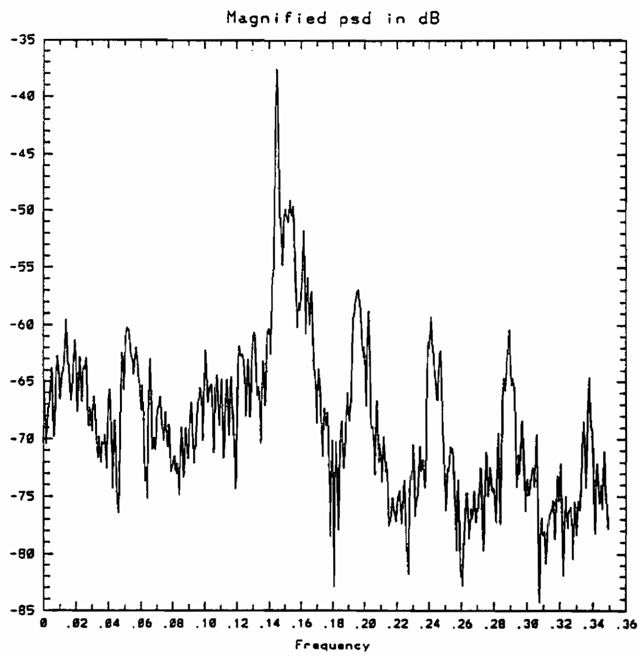


(b) Ion trajectories.

Figure 45: Simulation results at $\omega_o = 0.91$, with $ecenter = ngpsd = 181$ and Region II parameters.



(c) Power spectrum for the time series from $t = 400$ to $t = 2000$.



(d) Comprehensive power spectrum from zero frequency to 0.35. ($t = 400$ to $t = 2000$)

Figure 45 (continued).

harmonics (around 0.015 and 0.05) are also enhanced. In this case, the DM and possibly the UM[27] disappear from the power spectrum.

Similar spectra are obtained for the simulations at $\omega_o = 0.92$ and $\omega_o = 0.94$. In Figure 46(a), the BUM peaks at a frequency of 0.0065 above the pump. But the spectrum of the BUM does not stay unchanged when the other portion of time series is used for spectral calculation. It is generally very irregular. Its only constant property for most of the cases is that its bandwidth is broad. Also in Figure 46(a), a weak lower sideband appears, but it is not the DM because its frequency offset from the pump is smaller than a normal DM and it is not generated by the PDI involving lower hybrid waves. The energy enhancement at other frequencies is similar to the previous case (Figure 46(b)). The power spectrum for the case $\omega_o = 0.94$ is depicted in Figure 47(a). The BUM peaks at a frequency of 0.008 above the pump. Note that the settling time of the electric field time series increases in this case. The power enhancement at harmonics frequencies is more profound in Figure 47(b). Even the second cyclotron harmonic is enhanced.

When the pump frequency increases to $\omega_o = 0.96$, some different interactions occur in the plasma, although the BUM feature is still present in Figure 48(a). The electric field time series seems to collapse from a high amplitude to its steady state after $t = 500$. This phenomenon is typical for pump frequencies far away from the electron cyclotron harmonics. In this case, a sharp BUM peak occurs at a frequency of 0.0085 above the pump.

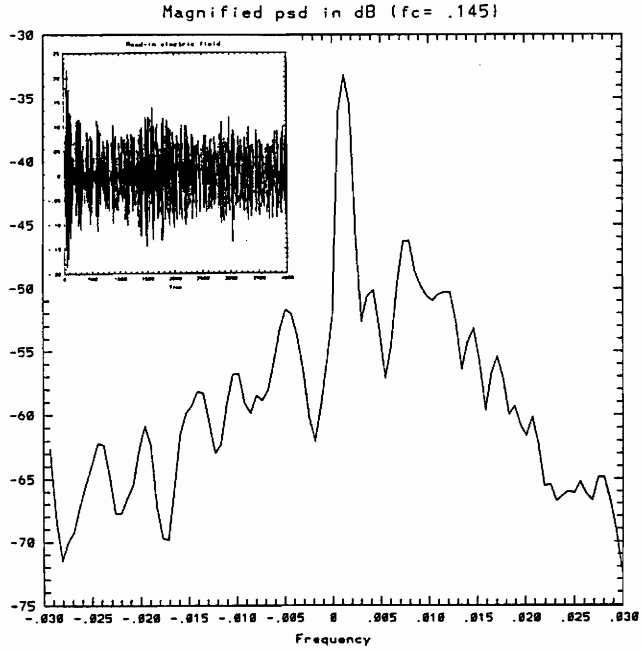
Ripple waves on the ion trajectories are found in the cases $\omega_o = 0.94$ and $\omega_o = 0.96$. But, they are relatively weak and are confined in the regions where the electron are heated. The ripple frequency is not constant and possibly, it somewhat relates to the broad bandwidth of the BUM. An illustration will be given later in the simulation results at $\omega_o = 1.22$ and $\omega_o = 1.26$.

If the pump frequency steps close to $3f_{ce}$, damping of the pump is expected. However, for $\omega_o = 0.905$, the BUM still appears in Figure 49(a). Its peak frequency is about 0.009 above the pump. In this case, the pump frequency is 5.5% above the third electron cyclotron harmonic. We may expect that an efficient mode conversion of the upper hybrid wave (the pump) into electron Bernstein modes, according to Rao and Kaup's theory (see section 3.8). The energy enhancement at higher cyclotron harmonics is large ($\sim 10dB$, similar to the reference case) in Figure 49(b). But, no significant power enhancement at lower harmonics. The DM feature is quenched because no PDI involving lower hybrid waves occurs.

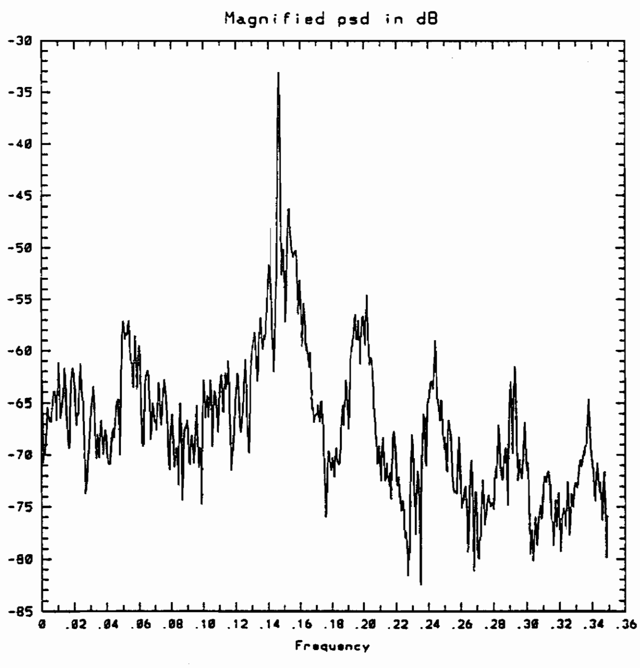
Fourth harmonic

When the pump frequencies step on to the fourth harmonic, the simulated spectra are in general

[27] It may be embedded in the suspected BUM structure.

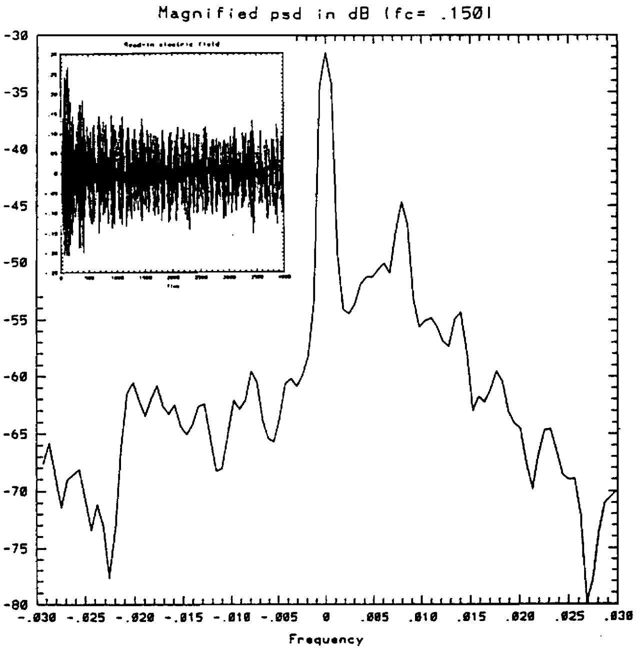


(a) Power spectrum for the time series from $t = 200$ to $t = 1800$.

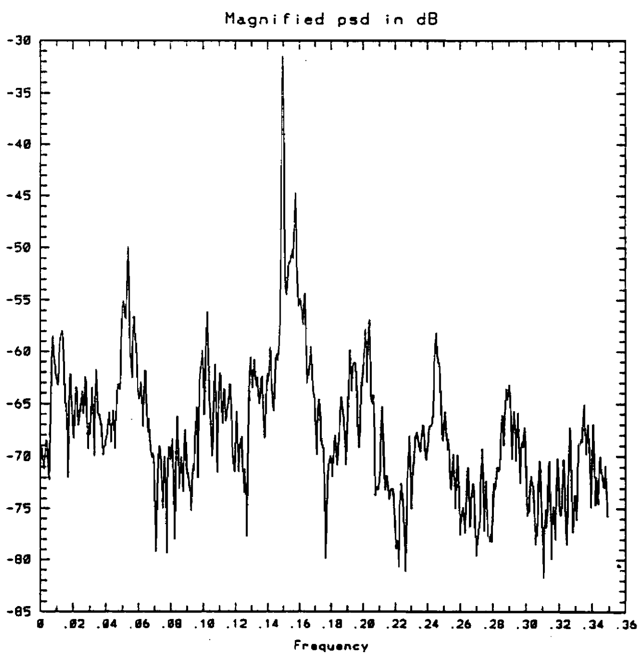


(b) Comprehensive power spectrum from zero frequency to 0.35. ($t = 200$ to $t = 1800$)

Figure 46: Simulation results at $\omega_o = 0.92$, with $ecenter = ngpsd = 190$ and Region II parameters.

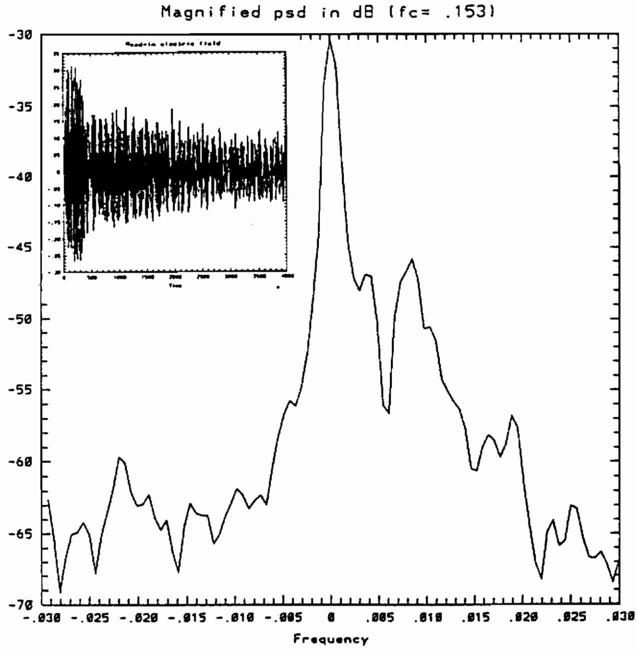


(a) Power spectrum for the time series from $t = 1000$ to $t = 2600$.

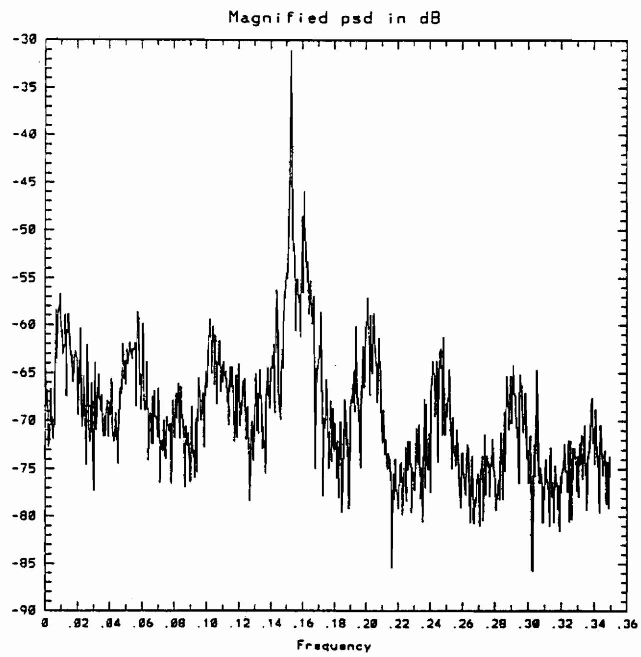


(b) Comprehensive power spectrum from zero frequency to 0.35. ($t = 1000$ to $t = 2600$)

Figure 47: Simulation results at $\omega_o = 0.94$, with $ecenter = ngpsd = 209$ and Region II parameters.

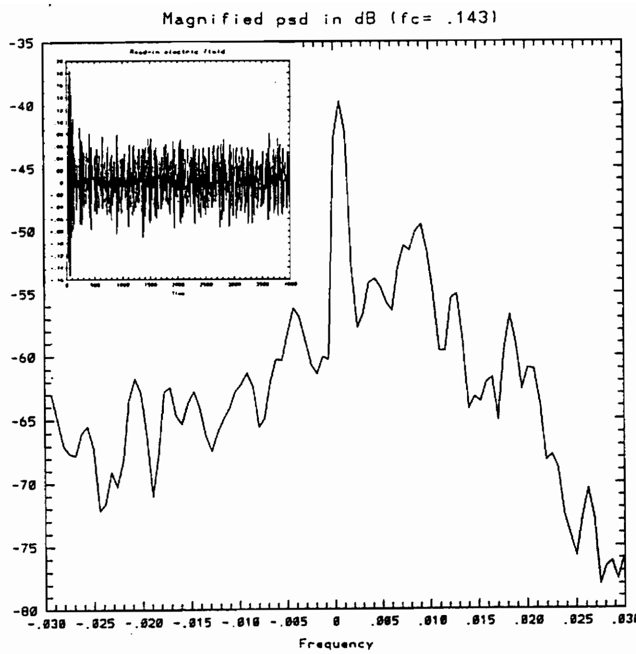


(a) Power spectrum for the time series from $t = 1000$ to $t = 2600$.

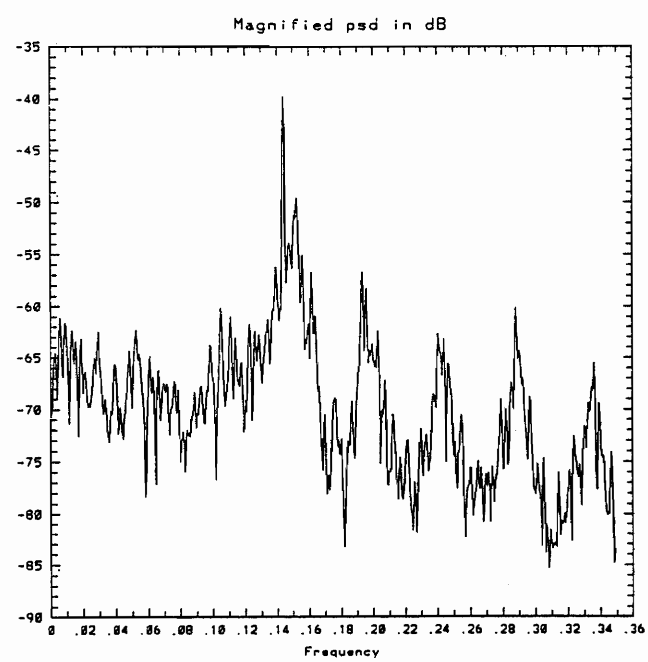


(b) Comprehensive power spectrum from zero frequency to 0.35. ($t = 400$ to $t = 3600$)

Figure 48: Simulation results at $\omega_o = 0.96$, with $ecenter = ngpsd = 228$ and Region II parameters.



(a) Power spectrum for the time series from $t = 1000$ to $t = 2600$.



(b) Comprehensive power spectrum from zero frequency to 0.35. ($t = 1000$ to $t = 2600$)

Figure 49: Simulation results at $\omega_o = 0.905$, with $ecenter = ngpsd = 176$ and Region II parameters.

very different from the cases at frequencies slightly above $3f_{ce}$. We do not spectilate until the discussion. Figure 50(a) to (d) show some important simulation results for the case $\omega_o = 1.22$. Region II parameters are used. The electron heating is localized around the pump region and tends to extend in negative x-direction (Figure 50(a)). Note that in Figure 50(b), there are some weak and irregular ripples superimposed on the ion trajectories below the pump region. The settling time of the electric field is similar to the case $\omega_o = 0.96$. However, the suspected BUM shrinks at least by half and shifts next to the pump in this case (Figure 50(c)). The energy levels near the sixth, the seventh and the eighth harmonics are enhanced in Figure 50(d).

Similar spectrum is obtained in Figure 51(a) for the case $\omega_o = 1.24$. Here, the sudden collapse of the electric field at $t = 1000$ is evident. Unlike the previous case, besides the power enhancement at higher harmonics, there is energy enhancement at lower harmonics (Figure 51(b)).

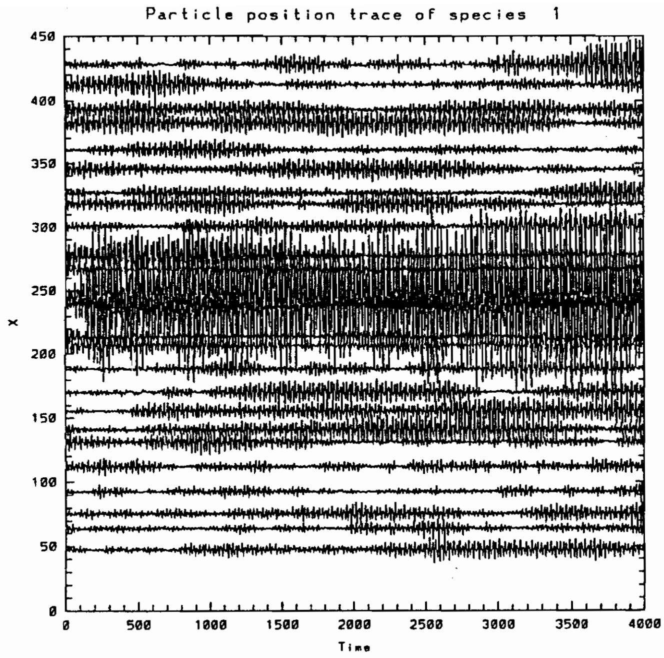
When the pump further increases to $\omega_o = 1.26$, the BUM merges with the pump in Figure 52(c). The electrons are heated in the pump region and below (Figure 52(a)). The reason of electron heating at the right edge of the plasma is unknown. Irregular ripples are present in the ion trajectories below the pump region (Figure 52(b)). The energy absorption of the pump is large and the electric field takes a long period to settle in this case. The power enhancement at higher harmonics is profound in Figure 52(d).

The BUM disappears when the pump reduces to $\omega_o = 1.205$. Instead, the DM and UM show up in Figure 53(a). But, in Figure 53(b), the low-frequency signature of the PDI involving lower hybrid waves cannot be found. Similar spectrum is obtained for the case $\omega_o = 1.21$ (Figure 54(a)). Here, the unknown feature centered at a frequency of 0.02 above the pump appears again. In both cases, there are power enhancements at higher harmonics. Some weak ripples are observed in the ion trajectories seemingly propagating in both directions from the pump region for the case $\omega_o = 1.21$.

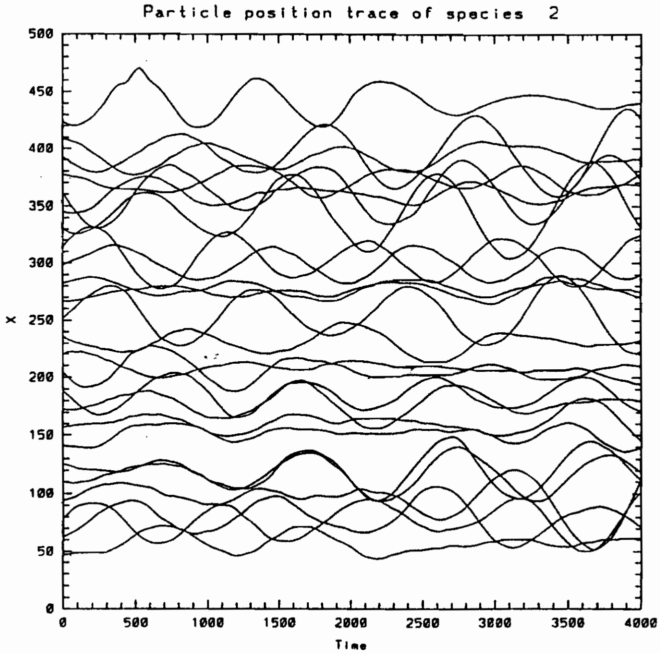
So far, the simulation results seem to point out that the DM and UM are associated with the parametric decay instability and the bidirectional propagation of lower hybrid waves in the form of regular ripples on the ion trajectories. On the other hand, the generation of the suspected BUM is related to electron heating in the pump region and below. In all cases, the development of SEE features depend on the excitation position in the plasma. We have shown the effect of pump position on the development of the DM and UM. In the following, two different excitation positions are simulated at $\omega = 1.205$ and $\omega_o = 1.22$. Their spectra are totally changed when compared with their previous cases.

Different pump positions

When the pump position is moved near the left edge of the plasma, the BUM appears in the spec-

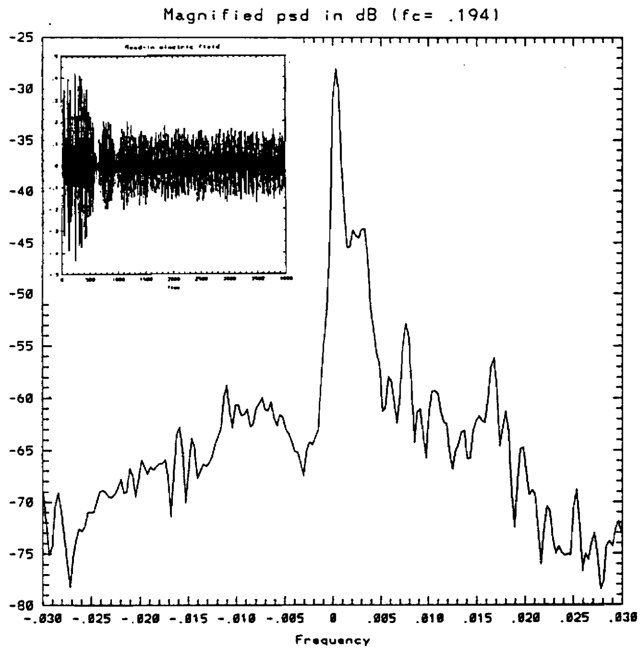


(a) Electron trajectories.

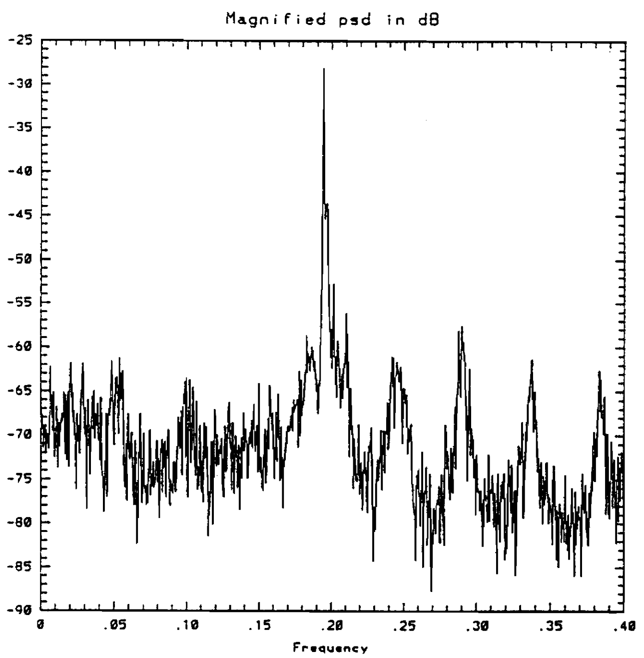


(b) Ion trajectories.

Figure 50: Simulation results at $\omega_o = 1.22$, with $ecenter = ngpsd = 511$ and Region II parameters.

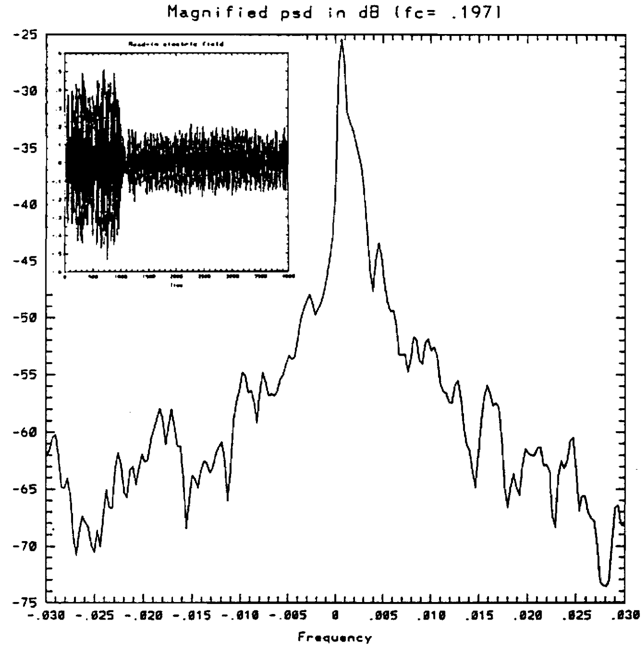


(c) Power spectrum for the time series from $t = 600$ to $t = 3800$.

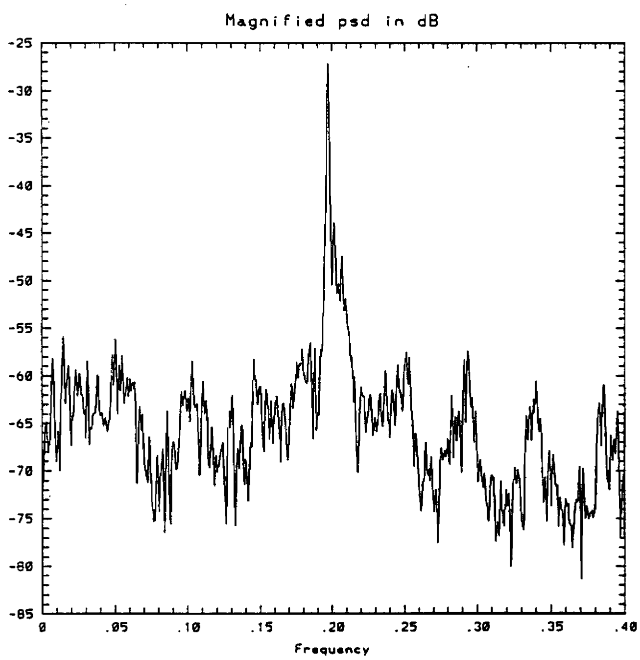


(d) Comprehensive power spectrum from zero frequency to 0.4. ($t = 600$ to $t = 3800$)

Figure 50 (continued).

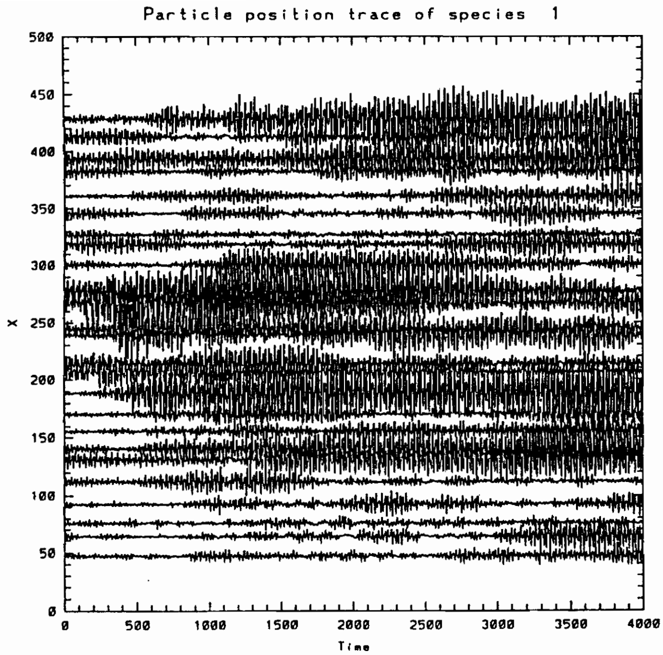


(a) Power spectrum for the time series from $t = 600$ to $t = 3800$.

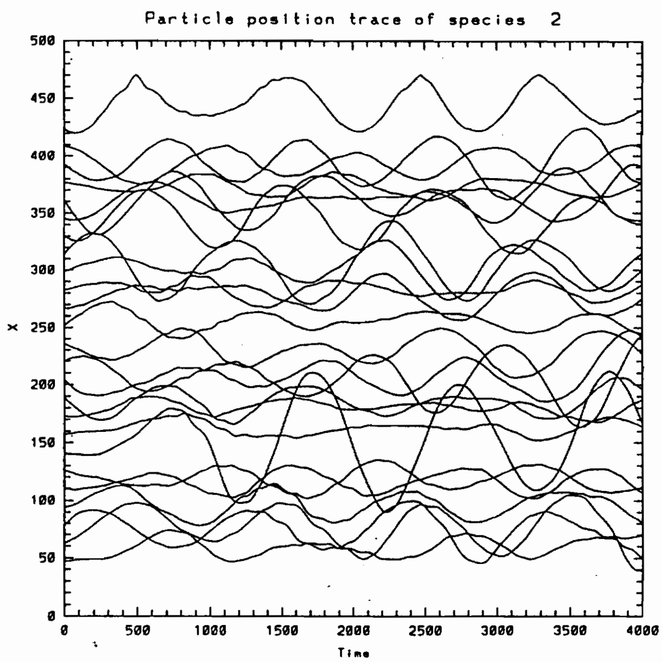


(b) Comprehensive power spectrum from zero frequency to 0.4. ($t = 1000$ to $t = 2600$)

Figure 51: Simulation results at $\omega_o = 1.24$, with $ecenter = ngpsd = 536$ and Region II parameters.

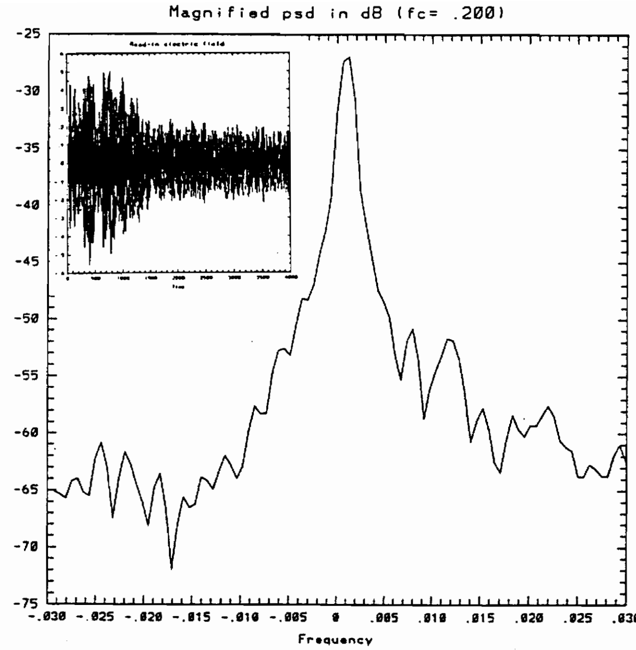


(a) Electron trajectories.

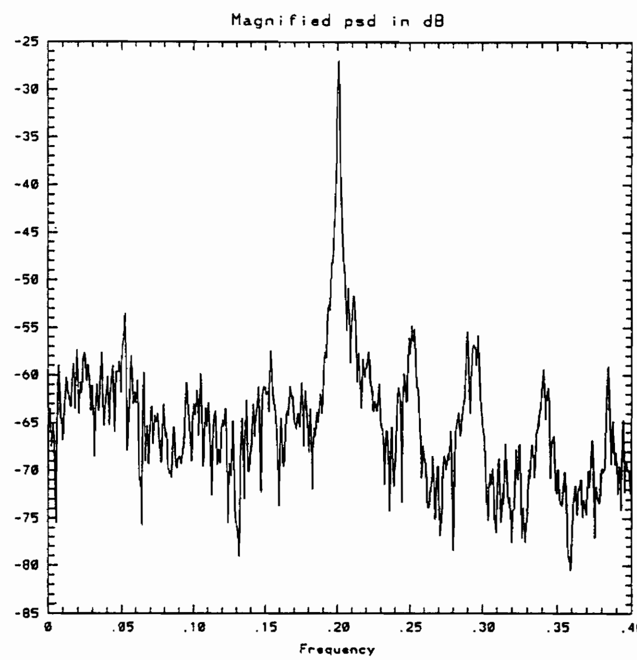


(b) Ion trajectories.

Figure 52: Simulation results at $\omega_o = 1.26$, with $ecenter = ngpsd = 561$ and Region II parameters.

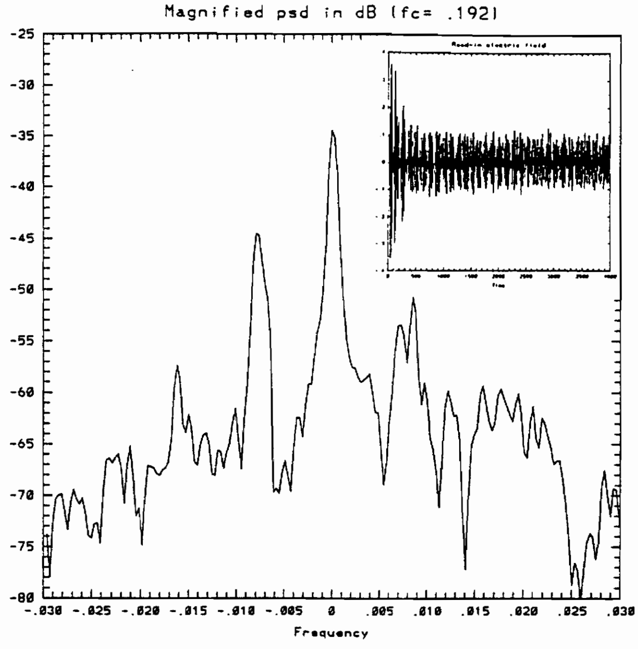


(c) Power spectrum for the time series from $t = 2000$ to $t = 3600$.

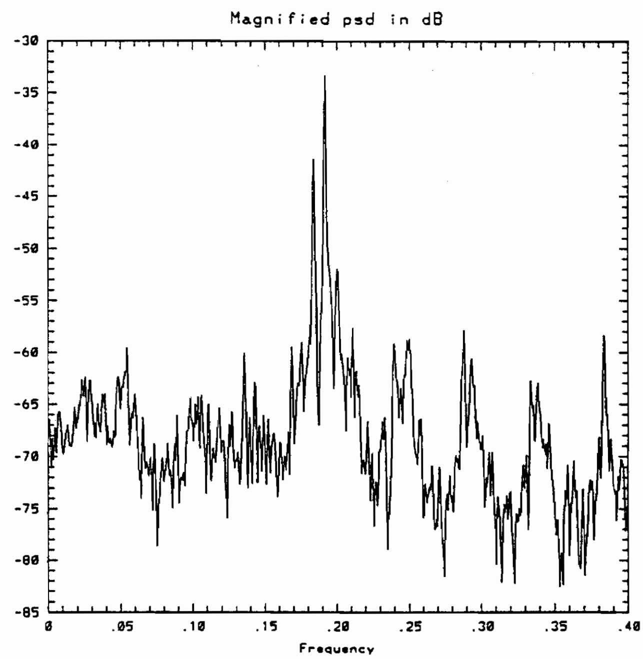


(d) Comprehensive power spectrum from zero frequency to 0.4. ($t = 2000$ to $t = 3600$)

Figure 52 (continued).

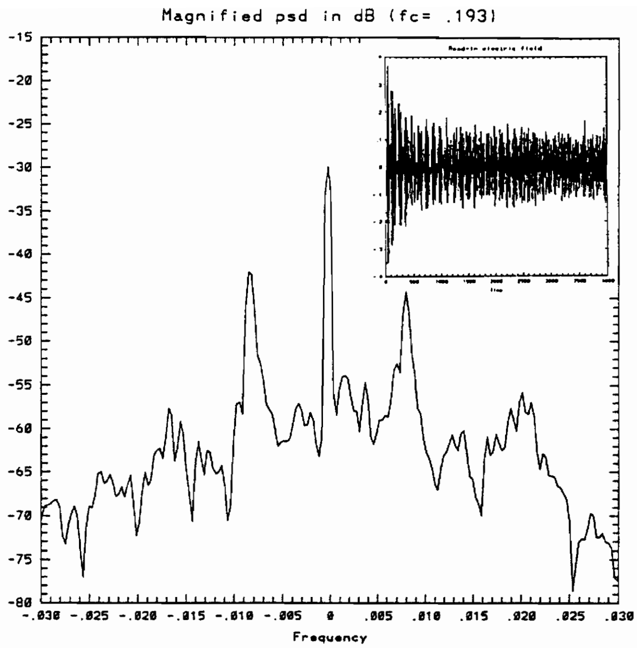


(a) Power spectrum for the time series from $t = 600$ to $t = 3800$.

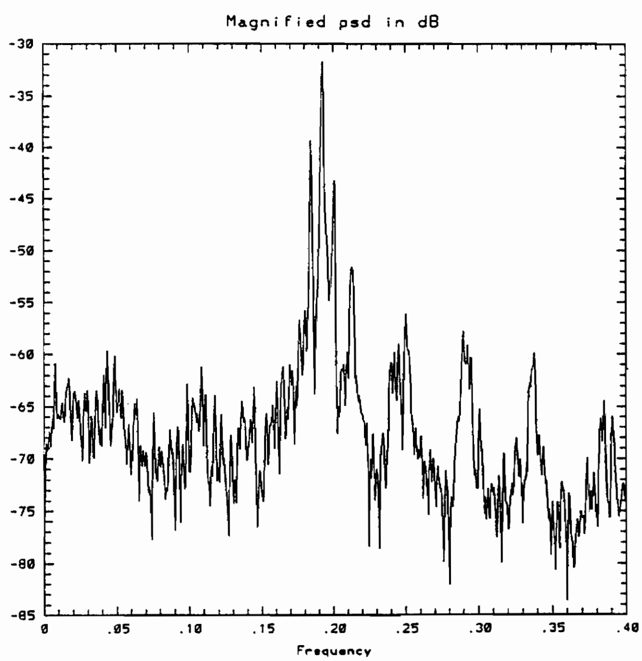


(b) Comprehensive power spectrum from zero frequency to 0.4. ($t = 1000$ to $t = 2600$)

Figure 53: Simulation results at $\omega_o = 1.205$, with $ecenter = ngpsd = 493$ and Region II parameters.

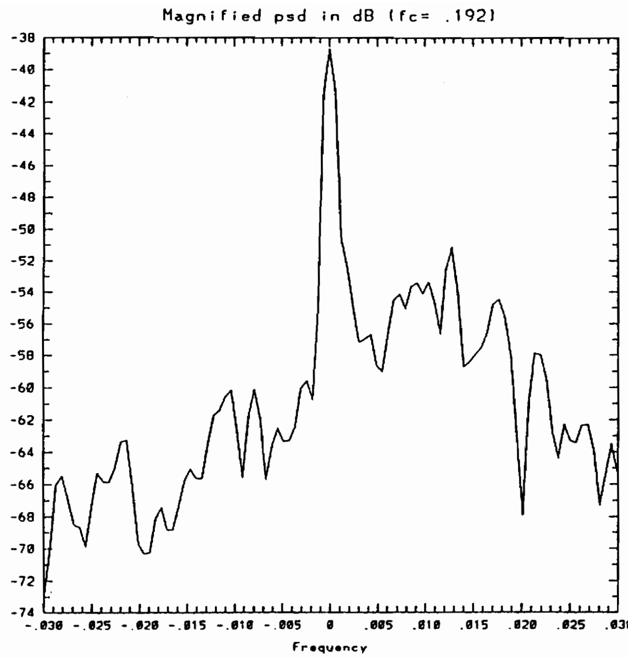


(a) Power spectrum for the time series from $t = 600$ to $t = 3800$.

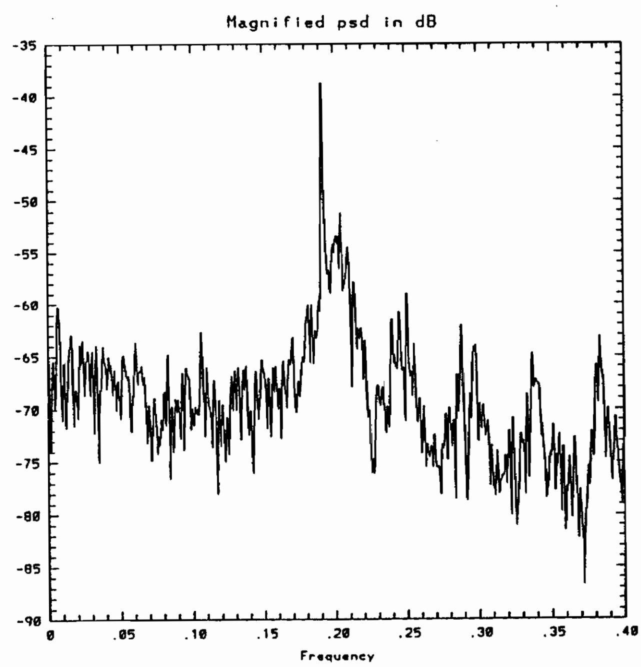


(b) Comprehensive power spectrum from zero frequency to 0.4. ($t = 1000$ to $t = 2600$)

Figure 54: Simulation results at $\omega_o = 1.21$, with $ecenter = ngpsd = 499$ and Region II parameters.



(a) Power spectrum. ($t = 1000$ to $t = 2600$)



(b) Comprehensive power spectrum from zero frequency to 0.4. ($t = 1000$ to $t = 2600$)

Figure 55: Simulation results at $\omega_o = 1.205$, with $ecenter = ngpsd = 177$ and other parameters similar to those shown in Figure 33(a).

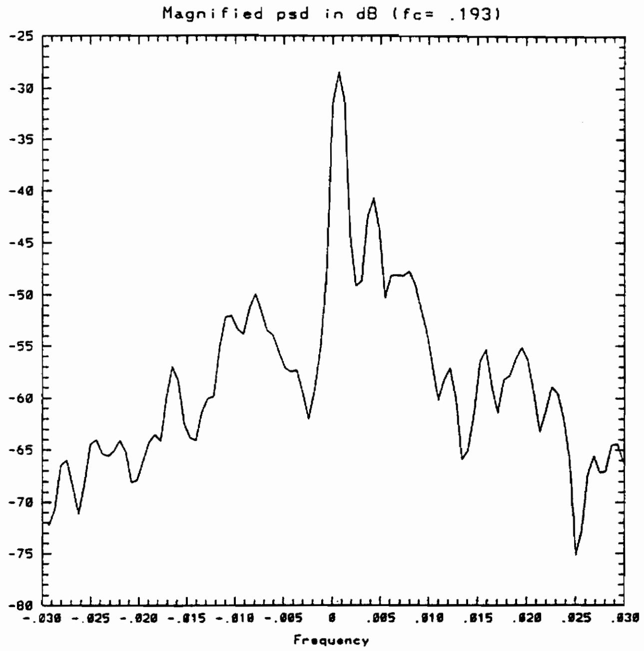
trum for the case $\omega_o = 1.205$ (Figure 55(a)). However, in Figure 55(b), there is a sign of the PDI at low frequency, and the power enhancement at higher harmonics are less profound. Figure 56 shows another case for the excitation position close to the right edge of the plasma. The pump frequency is at $\omega_o = 1.22$. The BUM is present at its nominal frequency with a frequency spike next to the pump.

Effects of mass ratio and electron and ion temperatures

Besides the pump positions, the mass ratio and the electron and ion temperatures may have effects on the SEE features for the pump frequency slightly higher than nf_{ce}. The following simulations are carried out to investigate these effects. Figure 57 depicts the power spectrum with $\omega_o = 0.91$ and $K = 225$. Two sidebands show up adjacent to the pump. The lower one is believed to be a weak DM, while the upper one is suspected to be the UM. The frequency of the BUM remains almost unchanged, but it splits up into two parts. In this case, it seems that the UM can be separated from the BUM and the simulated power spectrum resembles more closely the experimental data, where the mass ratio of the "real" electrons and O^+ ions is $1836 \times 16 = 29376$. If such a realistic mass ratio was used, the DM and UM would move 11.5 times closer than those in Figure 57, provided that the frequency resolution is sufficient to separate them out.

The effect of electron temperature on the BUM is shown in Figure 58(a) to (c) and Figure 59. First, the electron thermal velocity is decreased by one half. Consistent with the previous observation, the electron heating is localized in the pump region and below (Figure 58(a)). Some weak irregular ripples are found in the ion trajectories (Figure 58(b)). In Figure 58(c), the spectral characteristic of the BUM does not change much, but, a weak DM appears. Next, the electron thermal velocity is doubled. The resulting power spectrum is depicted in Figure 59. Again, the frequency and amplitude of the BUM do not change. For the effect of ion temperature, Figure 60 shows the power spectrum with a reduced ion thermal velocity. The amplitude of the BUM slightly increases, but its frequency remains about the same. From the above four numerical experiments, we may conclude that both electron and ion temperatures as well as the mass ratio have no significant effect on the BUM development. It seems that the physical process responsible for the generations of the DM and UM is fundamentally different from the BUM.

In most cases, some irregular waves on the ion trajectories coexist with the appearance of the BUM. The broad spectral width of the BUM may be caused by some dispersed interactions inside the plasma. Occasionally, deep fades occur within the BUM bandwidth. It may be due to insufficient data to densely fill up the whole BUM bandwidth. Thus, if the number of time steps increases, the simulated spectrum with the BUM feature may be more similar to the real spectrum.



Power spectrum. ($t = 1000$ to $t = 2600$)

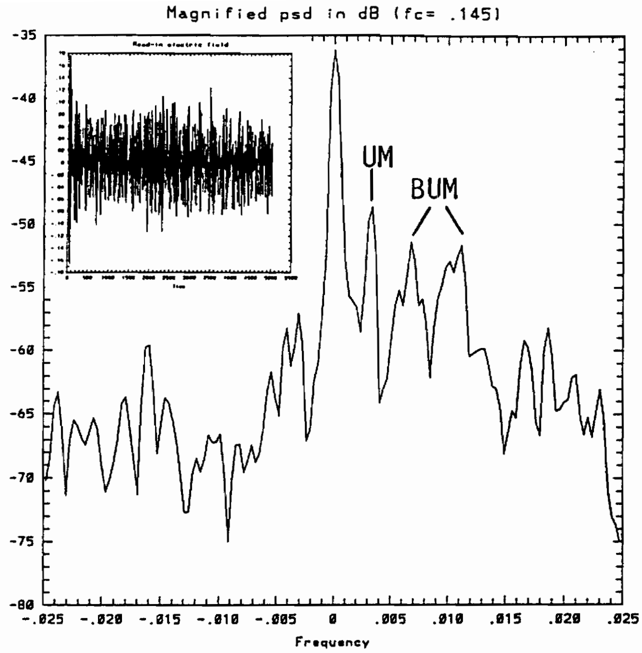
Figure 56: Simulation results at $\omega_o = 1.22$, with $ecenter = ngpsd = 773$ and other parameters similar to those shown in Figure 33(b).

5.3.4 Miscellaneous case

All the above discussions concern the spectral behaviors around the pump frequency when f_o is close to nf_{ce}. However, we may be interested in some other cases where the pump frequency is far away from the electron cyclotron harmonics. The following case corresponds to a commonly used heater frequency (5.1MHz) in ionospheric modification experiments. It corresponds to a simulation frequency of $\omega_o = 1.125$. The power spectrum is displayed in Figure 61(a). One cannot find the DM, UM and BUM. But, power enhancement occurs at some other frequencies in Figure 61(b). There is no significant power enhancement at electron cyclotron harmonics, except the zeroth and the fourth. The absorption of the pump energy is high in this case, which is a general fact for the pump wave far away from nf_{ce}.

5.4 Discussions

After performing a large number of numerical simulation experiments to study the SEE, we discuss some implications from the simulation results in this section. Since the 1-D plasma used in the



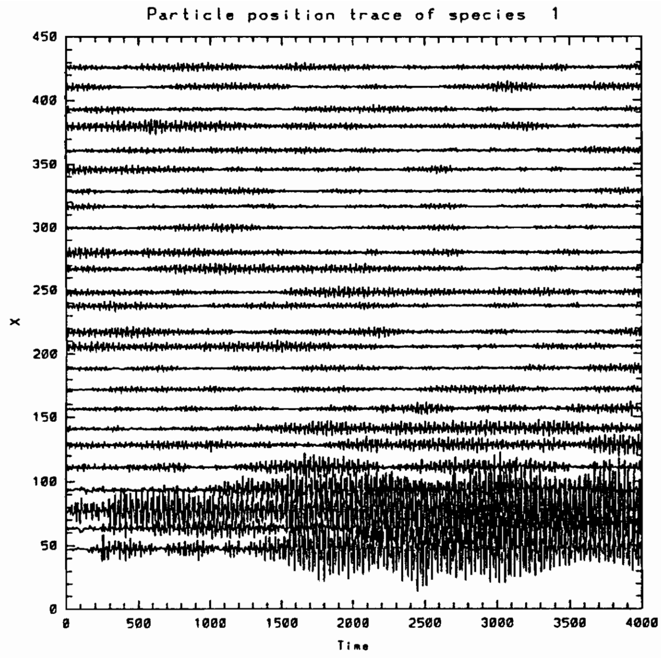
Power spectrum. ($t = 600$ to $t = 3800$)

Figure 57: Simulation results at $\omega_o = 0.91$, with $ecenter = ngpsd = 186$, $K = 225$ and Region II parameters.

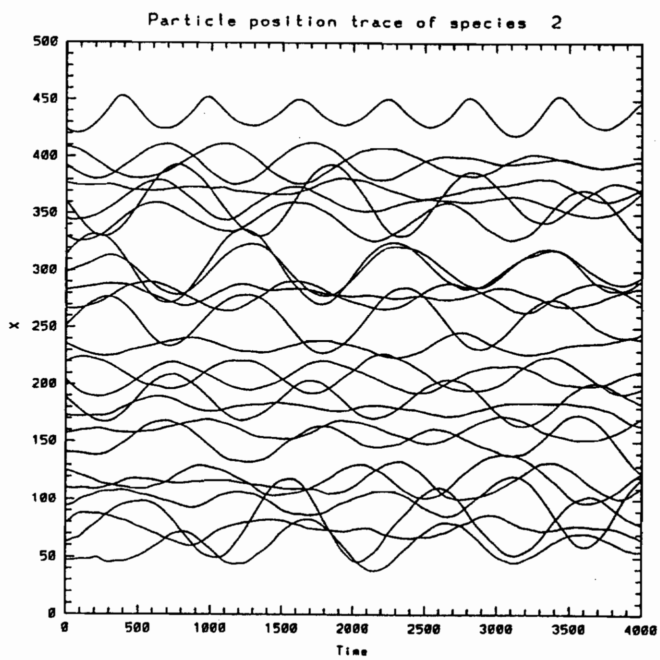
simulations only models the upper hybrid layer, interactions occurring at the reflection layer are not considered. Thus, we can only relate the simulation results to the discussions of the DM, UM, BUM and quenching of DM in the following. For the observed SEE features in the above power spectra, the appearance of the DM is very believable because we have a fairly well developed theory as a backup to characterize the DM. However, we need to be careful about the conclusions of the UM and BUM in the SEE simulation. The upper sidebands are presumed to correlate with these two features since simulations reproduce several characteristics of these features, but they are not 100% conclusive. Unless there is a contradiction, either from the experiments or the simulations, to state otherwise, we assume that they are the SEE features as demonstrated in the last section. Then we can proceed to the following summary about the observations from the SEE simulation.

1. The downshifted maximum

It is confirmed from the simulation that the DM is generated by the parametric decay instability involving the lower hybrid waves. The lower hybrid waves propagate in the form of regular ripples on the ion trajectories away from the pump region. The interaction of these propagating ion ripples

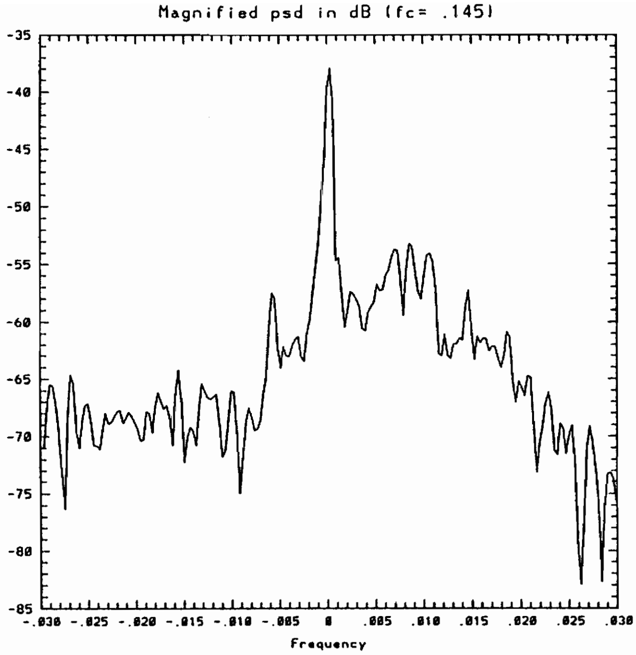


(a) Electron trajectories.



(b) Ion trajectories.

Figure 58: Simulation results at $\omega_o = 0.91$, with $ecenter = ngpsd = 181$, $v_{te} = 0.5$ and Region II parameters.



(c) Power spectrum for the time series from $t = 600$ to $t = 3800$.

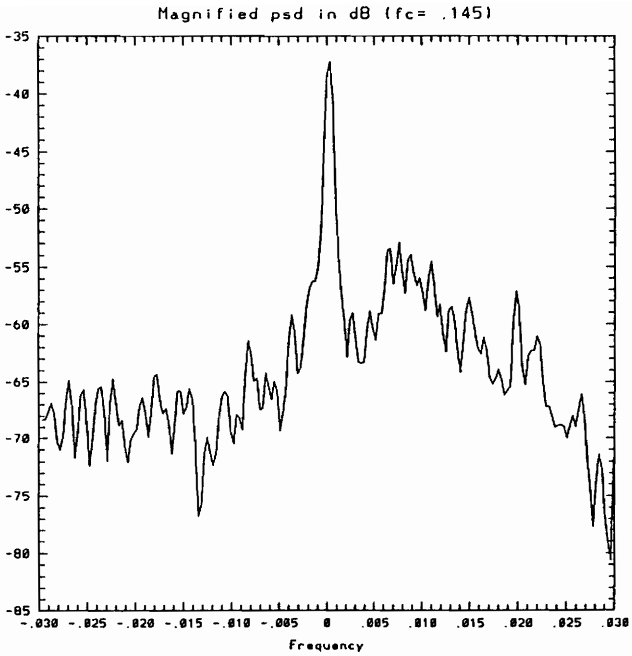
Figure 58 (continued).

and electrons can produce the DM spectrum all over the regions where the lower hybrid waves exist. Recall the dispersion of the lower hybrid waves from section 3.8,

$$\omega_l^2 = k_l^2 c_s^2 + \Omega_{LH}^2 \tag{175}$$

where $c_s^2 = (\gamma_e k_B T_e + \gamma_i k_B T_i)/m_i = \gamma_e v_{te}^2/K + \gamma_i v_{ti}^2$. With the simulation parameters in the reference case ($v_{te} = 1$, $v_{ti} = 0.0714$ and $K = 49$), the ion acoustic speed c_s is 0.277, assuming an adiabatic process for both γ_e and γ_i ($\gamma = 3$). We can estimate the wave number of the lower hybrid waves from some data obtained in the reference case. The lower hybrid frequency is 0.0064 and the frequency offset of the DM peak is about 0.01. Then, $k_l = 0.174$ or $\lambda_l = 2\pi/k_l = 36.1$. The wavelength is approximately equal to the one in the ion phase space at $t = 1400$ displayed in Appendix (C), where a solitary structure shows up in both electron and ion densities. Also from the simulation results, it seems that the PDI threshold is dominantly controlled by the ion temperature.

Using the above dispersion relation, it is straightforward to explain why the DM has a high-frequency cutoff at a lower hybrid frequency below the pump. The wave vector k_l is in general nonzero because of the wave number matching requirement of the PDI. Hence for the PDI involving lower hybrid waves, $\omega_l > \Omega_{LH}$. From the frequency matching of the PDI, $\omega_{DM} = \omega_o - \omega_l < \omega_o - \Omega_{LH}$.

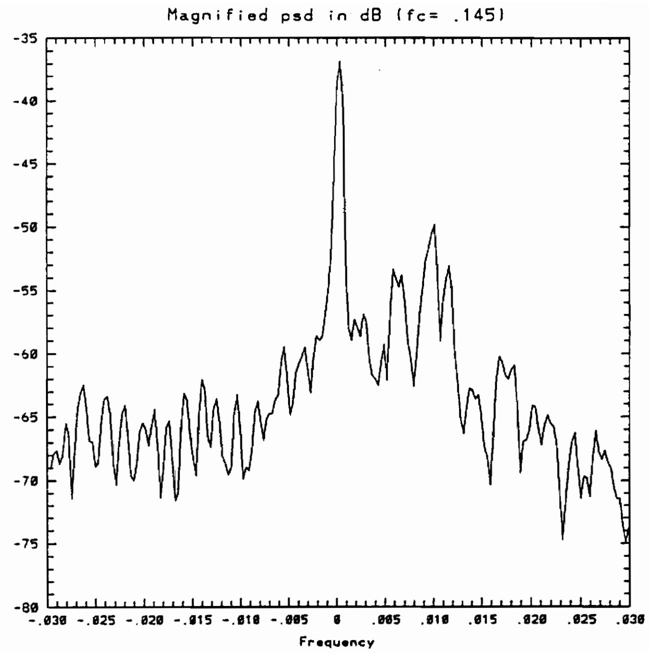


Power spectrum. ($t = 600$ to $t = 3800$)

Figure 59: Simulation results at $\omega_o = 0.91$, with *ecenter* = *ngpsd* = 181, $v_{te} = 2$ and Region II parameters.

Moreover, the lower hybrid frequency approximately depends on $1/\sqrt{K}$ (from (61)). When the mass ratio increases, the high-side cutoff frequency of the DM shifts towards the pump. This agrees with simulation results. The remaining question is how to set the wave number matching condition. In the SEE simulation, the pump wave (upper hybrid wave) parametrically decays into a lower hybrid wave and another plasma wave[28]. All the three wave numbers are nonzero. From the simulation results, we observed that the development of the DM and UM depends on the pump position. It may suggest that the wave numbers are determined by boundary conditions. We do not exactly know what these boundary conditions are and how they affect the development of the DM and UM. Theoretically, the DM can occur at all pump frequencies except at the electron cyclotron harmonics, whenever the wave number matching condition is satisfied. Nevertheless, these are difficult questions that we cannot provide answer for here because the interaction between the pump region and the boundaries is highly nonlinear. The propagation of the two decay waves is nonlinear. Furthermore, the boundary effects may be different from the 1-D plasma model used in the SEE simulation because at the two edges of the plasma, the propagating wave encounters a sharp change of density gradient

[28] This is different from the theory proposed by Leyser ([11] Leyser), which the third wave is an electromagnetic wave. We have already mentioned a difficulty in Leyser's theory about this electromagnetic wave in section 3.8.

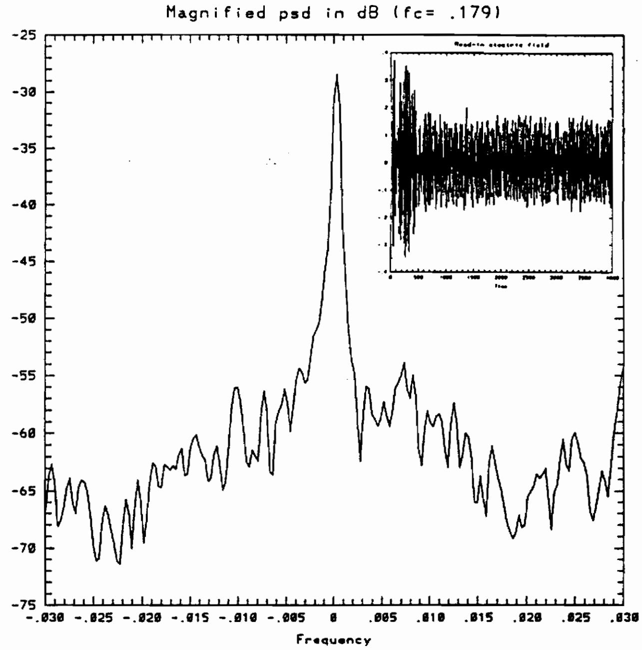


Power spectrum. ($t = 600$ to $t = 3800$)

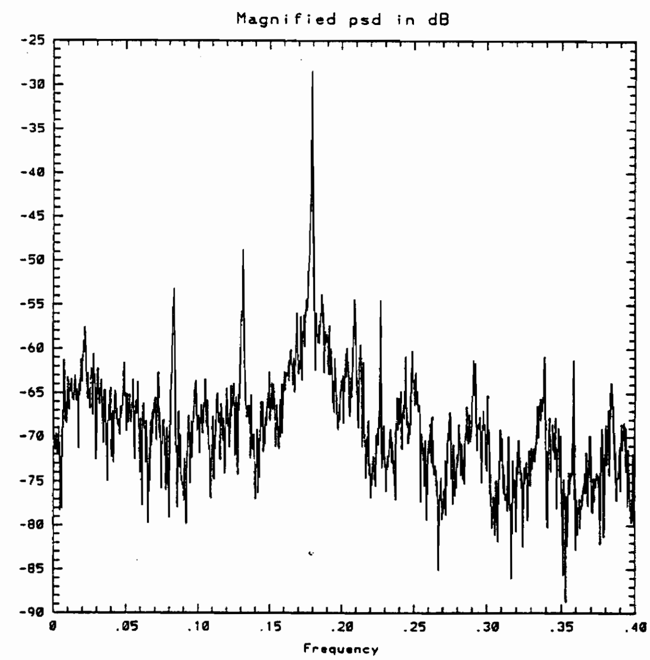
Figure 60: Simulation results at $\omega_o = 0.91$, with $ecenter = ngpsd = 181$, $v_{ti} = 0.0357$ and Region II parameters.

in the ionosphere, rather than a sharp cutoff of number of particles as in the simulation. A complete theoretical treatment may not be trivial.

The dependence of the offset frequency of the DM from the pump on electron and ion temperatures can also be understood using the dispersion relation. When the electron and/or ion temperature decreases, the ion acoustic speed is reduced. If we assume that k_l remains unchanged, the frequency of the lower hybrid wave (ω_l) will follow to decrease. On the other hand, when the electron and/or ion temperature increases, ω_l also increases. This agrees with the simulation results. But, from the fact that the decrease of the ion temperature causes more frequency shift than the electron temperature, it may indicate that γ_i is actually much higher than γ_e or the wave number k_l is changed by the ion temperature. The dependence of ω_l on v_{te} and v_{ti} may account for the empirical dependence of the frequency offset of the DM from the pump on the pump frequency noted by Stubbe ($\Delta f_{DM} \approx 2 \times 10^{-3} f_o$). As the pump frequency increases, the upper hybrid altitude and hence both electron and ion temperature also increase. Since the offset frequency Δf_{DM} is in fact the frequency of the lower hybrid wave, the DM shifts to lower frequencies as the ion acoustic speed is increased by the temperatures. The effect can be significant because the temperature changes rapidly with



(a) Power spectrum. ($t = 600$ to $t = 3800$)



(b) Comprehensive power spectrum from zero frequency to 0.4. ($t = 600$ to $t = 3800$)

Figure 61: Simulation results at $\omega_o = 1.125$, with $ecenter = ngpsd = 107$ and Region I parameters.

altitude in the F-region. However, the relation between Δf_{DM} and the pump frequency may not be directly proportional. The derivation has to consider the complicated temperature gradient of the ionosphere and the nonlinear relation between the ion acoustic speed and the temperatures in (168).

The nonlinearity of the interaction inside the plasma can be demonstrated by the fact that the propagation speed of the lower hybrid waves reduces as they propagate towards the ends of the plasma (as shown in several ion trajectories in the previous section and Appendix (C)). The group velocity of the lower hybrid wave is obtained by differentiating (168).

$$v_{gl} = \frac{d\omega_l}{dk_l} = \frac{c_s^2}{\omega_l/k_l} \tag{176}$$

Substituting the values calculated above, $v_{gl} = 0.212$. This roughly equals the propagation speed of the outgoing ripples immediately after leaving the pump region in the ion trajectories (Figure C-3(j)). But, it decreases rapidly as it approaches the ends of the plasma. The average outgoing speed of the lower hybrid waves is estimated to be 0.0925 in the last section. The reduction of the group velocity of the lower hybrid wave may be caused by the following reasons.

- The ratios of specified heat capacities (γ_e and γ_i) decrease as there is energy exchange between the lower hybrid waves and electrons.
- The formula for the group velocity which is derived from linear theory for a homogeneous plasma ([12] Nicholson) is not applicable because the plasma used in the SEE simulation is inhomogeneous and nonlinear.
- There is some dissipative effect inside the plasma that has not taken into account in the derivation of the dispersion relation.

Another observation from the simulation is that the lower hybrid wave seems to propagate more strongly towards the high density side. Again, this cannot be shown from the dispersion relation and the group velocity of the lower hybrid wave. In fact, many phenomena of the DM feature cannot be explained just using linear theory. Here are some of these unanswered puzzles observed from the simulation:

- What causes the bandwidth of the DM? It is not evident from the dispersion of the lower hybrid waves.
- How to account for the fact that the development of the DM is strong when the pump region is near the left edge of the plasma and when the pump frequency is close to but less than the second and the third electron cyclotron harmonics?

Besides the generation mechanism of the DM, the backward mode conversion raises questions from the simulation results. We have already discussed in section 3.8 that direct generation of an electromagnetic wave in the upper hybrid region is unfeasible. On the other hand, the DM frequency can be found when the observation point scans along the plasma around the pump region. In an occassion, the DM amplitude can be larger than the pump at a position far below the pump region. It is possible that the radiation point of the DM is far below its source point (upper hybrid point), and there may be more than one radiation points where the backward mode conversion occurs. Until now, we do not have a comprehensive theory of backward mode conversion in the ionosphere. Our discussion of the DM radiation has to stop at this stage. However, it needs to be noted that the backward mode conversion may act as a selective and filtering process for the spectral shape of the DM received on the ground. An electromagnetic code may be required to study this aspect.

2. The upshifted maximum

We presume that the UM is observed in the simulation because of the following reasons. It is always present with the DM and has an offset frequency from the pump less than the DM. In general, the UM is weaker than the DM although in several occassions, the reverse occurs. Most probably, its offset frequency from the pump depends on $1/\sqrt{K}$. The condition for its appearance is very similar to the DM. Therefore, we believe that the UM and DM are generated by the same process, namely the PDI involving lower hybrid waves. The relatively strengths of the UM and DM are determined by the other effects (e.g. boundary conditions) on the excitation position. However, unlike the DM, we do not have a reasonable account for the generation mechanism of the UM. A theoretical explanation for the UM is needed. It may be similar to the DM and can incorporate the above observation. In fact, many questions raised for the DM are also relevant for the UM. Whenever the DM and/or UM appear in the power spectrum, some power enhancements at electron cyclotron harmonics higher than the upper hybrid frequency are observed.

3. The broad upshifted maximum

We cannot completely be sure that the observed broad upper sideband is the BUM[29]. Nevertheless, it is presumed to be the BUM because of the following two reasons: it only appears when the pump frequency is slightly higher than nf_{ce}, and its bandwidth is broad. These are two of the

[29] A difficulty of the SEE simulation involving features other than the DM is that even when some plausible SEE features show up in the spectrum, we cannot confidently confirm that they are the structures which we search for. It is because until now, the most believable and reasonable SEE theory is the one concerning the DM, as discussed in section 3.8. If there is no theoretical justification for the intepretation of the simulation results, it may be easy to be fooled by the simulation results due to inappropriate parameter settings.

principal characteristics seen in the experiments. However, perhaps due to insufficent data points, it does not show a similar spectral shape as observed in the experiments. Besides the difference in spectral shapes, the simulation results do not agree with the empirical relation given by Leyser (section 2.4 and 3.8). An estimate for the frequency offset of the BUM peak above the pump can be calculated as follows.

$$\Delta f_{BUM} = f_{BUM} - f_o = f_o - nf_{ce} \tag{177}$$

For $n = 3$ and $\omega_o = 0.91$, $\Delta f_{BUM} = 0.0018$, and for $\omega_o = 0.92$, $\Delta f_{BUM} = 0.0034$. The observed simulated BUM spectra are located far away from these frequencies. In most cases, the BUM extends from a frequency offset of 0.006 to 0.013 above the pump. This frequency range remains relatively constant for all pump frequencies slightly above nf_{ce}. Some possible reasons to explain this discrepancy are :

- The broad upper sideband in the SEE simulation does not correspond to the BUM observed in the experiments. The BUM may not be generated in the upper hybrid layer.
- The simulation setup used in the previous results cannot reproduce the BUM feature.
- Some important dependence of the frequency of the BUM peak on other physical parameters is missed from the empirical relation.
- The empirical formula is in general not applicable to describe the relation of Δf_{BUM} and f_o. It roughly agrees with the experiments because of coincidence.

Since we do not even have a rough but confirmed explanation of the generation of the BUM, it has no way to determine the reason of discrepancy between the simulations and experiments.

Other observations of the simulated BUM are summarized as follows. They show nearly no dependence on the mass ratio and the electron and ion temperatures. Moreover, its generation does not involve the parametric decay instability or the lower hybrid waves. But, from the comparisons of Figure 53 and 55, and Figure 50 and 56, the development of BUM is sensitive to the pump position. Electron heating can be seen in the pump region and below. Occasionally, some weak and irregular ripples appear in the ion trajectories. It may indicate that the development of the BUM involves the interaction between waves and both electrons and ions. Also, power enhancements at nearly all electron cyclotron harmonics are observed when the BUM exists. Electron Bernstein modes may be triggered.

4. Quenching of DM

When the pump frequency is exactly equal to nf_{ce}, the pump wave is heavily damped. No or very weak PDI can occur, and the amplitudes of the DM and UM decrease rapidly or they are even quenched. Cyclotron damping seems to have been involved in such a case. The electron heating is only confined in the pump region. However, power enhancements at higher electron cyclotron harmonics are observed. Electron Bernstein modes may be triggered. From these observations, it is possible that both of the quenching mechanisms proposed by Leyser and Rao and Kaup can simultaneously exist. But, which mechanism dominates the damping cannot be determined from the simulations at present.

For Huang *et al*'s thermal OTSI proposal, it is impossible to verify using electrostatic plasma simulation. Probably, 1-D electromagnetic plasma simulation may be sufficient to investigate their theory.

5. Other SEE features

The SEE features which are not generated in the upper hybrid layer cannot be reproduced with the ES1G program. Hence, one cannot observe the continuum, DP and BSS in the simulated power spectrum[30]. As stated in section 3.8, these features are most likely generated in the reflection layer. They are left for the future study.

Under rare circumstances, the secondary SEE features such as $\frac{1}{2}DM$, $\frac{1}{2}UM$, $2DM$ and split DMs are observed in the simulated power spectra. Their developments are not clearly understood. But, some of them may be related to boundary effects.

In conclusion, the DM and UM are evidently generated by the same PDI involving the lower hybrid waves, while the generation mechanism of the BUM is still unknown. But, these two types of physical processes are very different in nature. They leave different signatures on the power spectra and electron and ion trajectories. Hopefully, the above new discoveries about the DM, UM and BUM from the SEE simulations can inspire the further theoretical developments of stimulated electromagnetic emissions.

[30] Even though the DP had appeared in the simulated spectrum, it could not be resolved with 8192 or 16384 data points.

6 Summary and conclusion

One-dimensional electrostatic plasma simulation using particle-in-cell technqiue has been performed to study some of the stimulated electromagnetic emission (SEE) features. Only the upper hybrid layer is modeled by the numerical simulation. Other features generated at the reflection layer are left the future work.

In our simulation results, mainly three features which we believe that they correspond to the downshifted maximum (DM), upshifted maximum (UM) and broad upshifted maximum (BUM) seen during SEE experiments are observed. It is also believed that these SEE features are generated in the upper hybrid layer. Amongst the three observed SEE features, the DM is best understood and investigated because it has a very believable theoretical explanation to account for its generation. It has been proposed by Leyser that the parametric decay instability (PDI) involving lower hybrid waves causes the DM feature. From the simulation, we confirmed that most likely, this is a valid conjecture. Several mass-ratio tests provide a striking evidence for the dependence of the high-frequency cutoff of the DM on the lower hybrid frequency and hence the existence of the PDI involving lower hybrid waves. However, detailed investigations of the generation process of the DM using the simulation raise a number of puzzles and difficulties in the oversimplified theory proposed for the DM. These are:

- The details of the mode conversion of the O-mode pump into the upper hybrid wave.
- The type of one of the daughter waves generated by the PDI at the DM frequency and the location(s) where the backward mode conversion occurs.
- Boundary effects on the wave number matching condition.
- A very limited frequency range around electron cyclotron harmonics for the existence of the DM.
- The bandwidth of the DM.
- The temperature (altitude) dependence of the DM development and structure.

On the other hand, due to insufficient theoretical foundations at present, the UM and BUM are somewhat more difficult to investigate using our present SEE simulations. Since a number of features observed in the experiments are reproduced by our simulations, we argue that the observed upper sidebands correspond to the UM and BUM. The frequency offset of the UM from the pump is less than the DM and the UM always coexists with the DM feature. This leads us to presume that the generation mechanism of the UM is probably related to the parametric instability involving lower hybrid waves. But, it is still not clear about its detailed generation process. It is suspected that the boundary effects on the wave number matching determine the relative strengths of the DM and UM. On the other hand, the BUM leaves a different signature on the power spectrum and the electron and ion trajectories from the DM and UM. Its bandwidth is broad and it appears only when the pump frequency is slightly above the electron cyclotron harmonics. However, the simulated spectrum does not conform to the empirical relation given by Leyser. A number of possible reasons has been suggested for this discrepancy.

The quenching of the DM is also observed in the simulation. It seems that both of cyclotron damping and mode conversion of upper hybrid waves into electron Bernstein modes take part in the DM quenching. However, it is difficult to determine which one dominates at present.

Concerning the future work of the SEE simulation, 1-D electromagnetic simulation is suggested to find out the detailed structure of the forward mode conversion and to directly verify the results simulated by the electrostatic code, as well as to study the backward mode conversion by monitoring the reflected electromagnetic waves. Moreover, a 2-D electrostatic or electromagnetic code may be used to investigate the interaction between the upper hybrid layer and the reflection layer, in order to provide a more comprehensive numerical study of the SEE features.

More theoretical analysis with the considerations of boundary effects and nonlinear wave propagation in an inhomogeneous plasma is required. Initial development can be based on simplified models[1] to study some primitive dependence of the DM structure on physical parameters (e.g. temperature effect on the PDI threshold and the shape of DM). The ultimate goal is to provide theoretical foundations for SEE simulations with more relevant ideas about what to search.

Regarding future experimental works, it is valuable to have chronological records of the SEE spectrum with ionograms at certain preselected frequencies over a day. It is because the spectral appearances of different kinds of SEE features at different known ionospheric conditions can provide information about the effects of density profile, temperatures, local density peaks[2], density cavities and irregularities on the development of SEE features. Certainly, it can facilitate theoretical developments as well as future SEE simulations to obtain a better understanding of SEE. Only when the underlying generation mechanisms of stimulated electromagnetic emissions are completely known, we can then use SEE as a diagnostic tool to investigate the ionosphere.

[1] For examples, the one used in this simulation work or a modified version with constant densities at both boundaries, which is similar to the density profile of the ionosphere.

[2] For examples, E and F1 peaks.

Appendix

(A) INIT subroutine listing

This is a FORTRAN subroutine which has been completely rewritten to implement the particle position and velocity loading described in section 4.2.4. It is called by the ES1G main program[1] during initialization. The density profile is linear and the velocity distribution is Maxwellian. Quiet start is used. The listing of the INIT subroutine is as follows.

```
cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
      subroutine init(il1,il2,m,q,t,nm,rho0,is)
c
c  loads particles one species at a time.
c
cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c
      common /param/ nsp, l, dx, dt, tb, sb, thetab
      real l,vv,fv,dv,df
      real lg, cdf(500000), dfn
      real m, nm, xmax, xmin, lm, xo, nlm
      common x(500000),vx(500000),vy(500000),vz(500000)
      common /rnorm/ vthrm(5),no(4),dfn(4),aa,bb,cc
      common /xmap/ ixmap(125000,4),n(4)
      real no
      data twopi/6.2831 85307 17958/
c
c  lm = length with loaded particles. (xmax-xmin)
c  xo = midpoint of plasma length (xmin+xmax)/2
c  cdf(i) = cumulative frequency for v.
c  fv   = distribution function for v.
c  ixmap(i) = Index (i) represents the current exchanged index of the ith
c        particle of the original loading. (used in particle tracing)
c
c
c  INPUT variables:
c
c  n(is)  =number of particles (for each species).
c  wp     =plasma frequency.
c  wc     =cyclotron frequency.
c  qm     =q/m charge:mass ratio.
c  vt1    =rms thermal velocity for random velocities.
c  nlg    =number of loading groups (sharing same ordered velocities).
c  nv2    =multiply maxwellian (for ordered velocities) by v**nv2.
c  v0     =drift velocity.
c  xmin   =start position of the plasma (Normalized to 1 when input)
c  xmax   =end position of the plasma (Normalized to 1 when input)
c  dfn   =total difference in electron density (normalized to mean density no)
c  mode, x1, v1, thetax and thetay are for loading a sinusoidal perturbation.
c  velocity contributions thru vt1, vt2, v0 and v1 are additive.
```

[1] A complete listing of the program ES1G is available from Dr. W.A. Scales, Room 615, Whittemore Hall, Department of Electrical Engineering, Virginia Tech, Blacksburg, VA 24061.

```
c
c  default input parameters:
c
      data wp,wc,qm,vt1,nlg,nv2,v0,mode,x1,v1,thetax,thetav,
     . xmin,xmax /1.,0.,-1.,0.,1,0,0.,1,.001,4*0.,1./

c
c     read species input
c

      read(17,*)n(is),wp,wc,qm,vt1,v0,mode,x1,v1,xmin,xmax,dfn(is)

c
c
      vthrm(is)= vt1
      t=tan(-wc*dt/2.)
      il2=il1+n(is)
c
      if (xmin.lt.0.) then
        write(*,*) 'XMIN out of range.'
        xmin=0.
      endif
      if (xmax.gt.1.) then
        write(*,*) 'XMAX out of range.'
        xmax=1.
      endif
c
      xmin=xmin*l
      xmax=xmax*l
      lm=xmax-xmin
      q=lm*wp*wp/(n(is)*qm)
      m=q/qm
      nm=n(is)*m

      xo=0.5*(xmin+xmax)
      no(is)=1.0*n(is)/lm
      dfn(is)=dfn(is)*no(is)
      if (abs(dfn(is)).gt.(2.*no(is))) then
        write(*,*) 'Too big DFN.'
        dfn(is)=sign((2.*no(is)),dfn(is))
      endif
      if (lm.lt.0.) then
        write(*,*) 'Negative plasma length.'
        xmin=0.
        xmax=l
        lm=l
      endif
      if (is.eq.nsp) then
        aa=0.5*dfn(nsp)/lm
        bb=no(nsp)-dfn(nsp)*xo/lm
        cc=(0.5*dfn(nsp)*xmax-n(nsp)*1.)*xmin/lm
      endif
c
c    Position loading according to the specified density gradient
c
      if (dfn(is).ne.0.0) then
c
c  Use quadratic solution for x(i).
c
        b=n(is)-dfn(is)*xo
       c1=(dfn(is)*xmax/2-n(is))*xmin
       do 100 i=1,n(is)
```

```
                i1=i-1+il1
                c=c1-i*lm
                x(i1)=(sqrt(b*b-2*dfn(is)*c)-b)/dfn(is)
 100   continue
       else
c
c  Uniform density
c
        ddx=lm/n(is)
        do 112 i=1,n(is)
                i1=i-1+il1
                x(i1)=(i-0.5)*ddx+xmin
 112    continue
      endif

*
*                   Load particles in 3D velocity space
*
c
c               Nondrifting Maxwellian velocity distribution
c
c  first get indefinite integral of cumulative distribution function
c  use midpoint rule - simple and quite accurate.
c
      do 70 i=1,n(is)
        ixmap(i,is)=i
 70   continue
      if (vt1.eq.0.) then
         i1=il1
         j=1
         do 60 i=1,n(is)
            vx(i1)=0.
            vy(i1)=0.
            vz(i1)=0.
            i1=i1+1
 60      continue
         goto 61
      endif
      vmax=5.*vt1
      dv=2.*vmax/(n(is)-1)
      cdf(1)=0.
      do 30 i=2,n(is)
         vv = ((i-1.5)*dv-vmax)/vt1
         fv = exp( -.5*vv**2 )
         cdf(i) = cdf(i-1)+amax1(fv,0.)
   30 continue
c
c  for evenly spaced (half-integer multiples) values of the integral,
c  find corresponding velocity by inverse linear interpolation.
c                     (vx = vy = vz)
c
      df=cdf(n(is))/n(is)
      i1=il1
      j=1
      do 40 i=1,n(is)
         fv=(i-.5)*df
   41    if(fv.lt.cdf(j+1)) go to 42
         j=j+1
         if(j.gt.(n(is)-1)) go to 45
         go to 41
   42    vv=dv*( j-1+(fv-cdf(j))/(cdf(j+1)-cdf(j)) )-vmax
         vx(i1) = vv
```

```
        vy(i1) = vv
        vz(i1) = vv
        i1=i1+1
   40 continue
   45 continue
c
c     Randomize particle positions and velocities by random
c     pair exchange to decorrelate phase space
c
      do 50 i=1,n(is)
      i1=i-1+il1
c
      ii = (n(is)-1)*ran1(idum) + il1
      xx = x(i1)
      x(i1) = x(ii)
      x(ii) = xx
      ixtemp=ixmap(i,is)
      ixmap(i,is)=ixmap((ii-il1+1),is)
      ixmap((ii-il1+1),is)=ixtemp
c
      ii = (n(is)-1)*ran1(idum) + il1
      vxx = vx(i1)
      vx(i1) = vx(ii)
      vx(ii) = vxx
c
      ii = (n(is)-1)*ran1(idum) + il1
      vyy = vy(i1)
      vy(i1) = vy(ii)
      vy(ii) = vyy
c
      ii = (n(is)-1)*ran1(idum) + il1
      vzz = vz(i1)
      vz(i1) = vz(ii)
      vz(ii) = vzz
   50 continue
 61   continue
c
c  Add perturbation
c  loading x(t=0), v(t=0), remember so no dt/2 correction.
c
      do 91 i=1,n(is)
        i1=i-1+il1
        theta=twopi*mode*x(i1)/l
        x(i1)=x(i1)+x1*cos(theta+thetax)
        vx(i1)=vx(i1)+v1*sin(theta+thetav)
 91   continue
c
c  apply boundary conditions, collect charge density, etc.
c
      call setrho(il1,il2-1,q,n(is)*q/lm)
      return
      end
```

(B) PSD program listing

This is a postprocessing program for the ES1G program. It reads the output file ES1G.DAT generated by the ES1G program and computes the power spectrum. A variety of selections are provided by the PSD program. These include the time series portion for the calculation of power spectrum, three options of filtering or smoothing the input/output data, data windows, and three different magnifications of the output power spectrum. An example of how to use the PSD program can be found in Appendix (C). Here is the listing of the PSD program.

```
c *************************
c * program psd.f (Ver. 6)
c *************************
c
c  By  K.T. Cheng
c  (28 Mar 1993)
c  (29 Mar 1993  revised)
c  (4 Apr 1993 2nd revised)
c
c Calculation of power spectrum using FFT.
c
c ********************************************************************
c Declaration of variables
c
        common pi
        common/chkrng/irs,dbmin,dbmax
        real pi,f,edata(65536),psd(65536),data1(131072),win(65536)
        real xe(65536),xpsd(65536),dt,en,n1,t0,t1,t2,edata1,df,ftemp
        real psddb(65536),mpsddb(65536),fc,deltaf,f1,f2,fmax,fmin
        real emax,emin,efft(131072),xp(65536),mpsd(65536),mr,fc1
        real xe1(65536),etemp,psdtemp,psdavg,eavg,dbmin,dbmax,dbtemp
        integer j,jj,k,i,dw,m,nt,ngpsd,N,ip,ek,j1,j2,nmax,nmin
        integer ir,i1,i2,npl,ms,np2,if1,if2,ideltaf,k1
        integer idm(3),is,rs,irs
        logical data_ex
        character sp*56
        character*40 mlabel1,mlabel2
        pi=3.141592654
c ********************************************************************
c
c Check es1g.dat
c
        inquire(FILE='es1g.dat', EXIST=data_ex)
        if (data_ex.eqv..true.) goto 500
        write(*,*) 'es1g.dat not found.'
        stop
c
c Initialize GKS
c
 500    continue
        call OPNGKS
c
c Read edata, compute emax & emin
c
        open(UNIT=8, FILE='es1g.dat', STATUS='old')
```

```
      read(8,*) dt,nt,ngpsd
      if (nt.gt.32768) then
              write(*,*) 'Too many data points!'
              stop
      endif
      write(*,140) nt,dt
 140  format('No. of read-in data points=',i5,', Time-step=',e10.3)
      do 301 i=1,nt
              read(8,*) edata(i)
              if (i.eq.1) then
                      emax=edata(i)
                      nmax=i
                      emin=edata(i)
                      nmin=i
                      goto 301
              else
                      if (edata(i).gt.emax) then
                              emax=edata(i)
                              nmax=i
                      endif
                      if (edata(i).lt.emin) then
                              emin=edata(i)
                              nmin=i
                      endif
              endif
 301  continue
      close(8)
c ********************************************************************
c
      write(*,*)'Select type(s) of data manipulation :'
      write(*,*)'(1) Smooth read-in electric field.'
      write(*,*)'(2) Smooth computed fft spectrum.'
      write(*,*)
     .'(3) Double no. of data points by linear intepolation'
      write(*,*)'(4) Do (1) & (2).'
      write(*,*)'(5) Do (1) & (3).'
      write(*,*)'(6) Do (2) & (3).'
      write(*,*)'(7) Do (1),(2) & (3).'
      write(*,*)'(8) Do none of above.'
      read(*,*) is
      if ((is.lt.1).or.(is.gt.8)) then
              write(*,*)'Wrong selection!'
              stop
      endif
      idm(1)=0
      idm(2)=0
      idm(3)=0
c
      go to (310,310,310,340,350,360,370,380), is
 310  idm(is)=1
      goto 380
 340  idm(1)=1
      idm(2)=1
      goto 380
 350  idm(1)=1
      idm(3)=1
      goto 380
 360  idm(2)=1
      idm(3)=1
      goto 380
 370  idm(1)=1
      idm(2)=1
```

```
      idm(3)=1
 380  continue
c *******************************************************************
c
c  Double no. of data points by linear interpolation.
c
      if (idm(3).eq.1) then
       do 801 i=1,nt
              k=nt+1-i
              edata(k+k-1)=edata(k)
 801   continue
       do 802 i=1,nt-1
              edata(i+i)=0.5*(edata(i+i-1)+edata(i+i+1))
 802   continue
       edata(nt+nt)=0.5*edata(nt+nt-1)
       dt=dt/2
       nt=2*nt
      endif
c *******************************************************************
c
c  Smooth read-in electric field.
c
      if (idm(1).eq.1) then
       etemp=edata(1)
       do 810 i=1,nt-1
              eavg=0.25*etemp+0.5*edata(i)+0.25*edata(i+1)
               etemp=edata(i)
              edata(i)=eavg
 810   continue
       edata(nt)=0.25*etemp+0.75*edata(nt)
      endif
c *******************************************************************
c
c Display edata and select data segment for fft.
c
      col=56
      row=20
      n1=nt/(row-1)
      en=(emax-emin)/(col-1)
      sp='                                                        '
      write(*,200) emax,(nmax-1)*dt,emin,(nmin-1)*dt
 200  format(3x,'TIME',6x,'EDATA',3x,'(emax=',e10.3,' @',e8.2,', emin=',
     .             e10.3,' @',e8.2,' )')
      do 210 i=1,row
              if (i.eq.1) then
                      ip=1
              else
                      ip=nint((i*1.0-1)*n1)
              endif
              t0=(ip-1)*dt
              edata1=edata(ip)
              ek=nint((edata1-emin)/en)
              write(*,220) t0,edata1,sp(:ek)
 220          format(e10.2,' ',e10.2,'| ',a,'*')
 210  continue
      write(*,*)'Data segment for FFT (must be more than 1 cycle)'
      write(*,*)'Input start time, end time :'
      read(*,*) t1,t2
c
c t1,t2 need not be exact integral numbers of dt because they are
c rounded off to their nearest integral numbers of dt in the following.
c For the whole data bank, enter
```

```
c        0,<any number greater than the displayed end time>.
c
        if (t1.lt.0.0) t1=0.0
        if (t2.gt.(nt-1)*dt) t2=(nt-1)*dt
        j1=nint(t1/dt+1)
        j2=nint(t2/dt+1)

        nj=j2-j1+1
        N=int(log((nj)*1.0)/log(2.0))
        if (nj.eq.nint(2**(N*1.0))) then
                jj=nj
                goto 150
        endif
        jj=int(2**(N*1.0+1.0))
 150    write(*,*)
        write(*,*) 'Total no. of data points selected = ',nj
        write(*,*) 'Total no. of data points for FFT = ', jj
        write(*,*)
     .'Total no. of read-in/interpolated data points = ',nt
        write(*,*) 'Time step = ',dt
        write(*,*) 'Grid cell under analysis = ', ngpsd
        write(*,*)
c ********************************************************************
c
c Zero patching
c
        do 52 i=1,jj
                if (i.gt.nj) then
                        efft(i)=0.0
                else
                        efft(i)=edata(j1+i-1)
                endif
52      continue

        jj2=2*jj
c
c Data windowing
c
        write(*,*) 'Data windowing (select a number):'
        write(*,*) '(1) square'
        write(*,*) '(2) Welch (parabolic)'
        write(*,*) '(3) Hanning (raised cosine)'
        write(*,*) '(4) Parzen (linear)'
        read(*,*) dw

        wss=0
        if (dw.eq.1) then
                do 501 m=1,jj
                win(m)=1.0
 501            continue
                goto 100
        elseif (dw.eq.2) then
                do 502 m=1,jj
                win(m)=1.0-(((m*1.0-1)-0.5*(jj*1.0-1))
     .                 /(0.5*(jj*1.0+1)))**2
 502            continue
                goto 100
        elseif (dw.eq.3) then
                do 503 m=1,jj
                win(m)=0.5*(1-cos(2*pi*(m*1.0-1)/(jj*1.0-1)))
 503            continue
                goto 100
```

```
      elseif (dw.eq.4) then
              do 504 m=1,jj
              win(m)=1.0-abs(((m*1.0-1)-0.5*(jj*1.0-1))
     .                         /(0.5*(jj*1.0+1)))
 504          continue
              goto 100
      else
              write(*,*) 'Wrong number!'
              stop
      endif
 100  continue
c ***********************************************************************
c
c Range selection for psd plots in dB.
c
      write(*,*)'Select dB range of psd plots :'
      write(*,*)'(1) Auto range.'
      write(*,*)'(2) Default range (-10,-70).'
      write(*,*)'(3) Enter range.'
      read(*,*) rs
      if ((rs.lt.1).or.(rs.gt.3)) then
        write(*,*)'Wrong selection'
        stop
      endif
      goto (510,520,530), rs
 510  irs=0
      goto 540
 520  irs=1
      dbmin=-70
      dbmax=-10
      goto 540
 530  write(*,*)'Enter min and max values of dB range :'
      read(*,*) dbmin, dbmax
      if (dbmin.gt.dbmax) then
        dbtemp=dbmax
        dbmax=dbmin
        dbmin=dbtemp
      endif
      irs=1
 540  continue
c ***********************************************************************
c
c Prepare data1 for fft.
c
      do 101 i=1,jj
              wss=wss+win(i)*win(i)
              data1(i+i-1)=efft(i)*win(i)
              data1(i+i)=0.0
 101  continue
      wss=jj*wss
      call four1(data1,jj,1)
c
c Compute psd
c
      do 102 i=1,jj/2+1
              k=jj/2+i-1
              psd(k)=(data1(i+i-1)*data1(i+i-1)
     .                +data1(i+i)*data1(i+i))/wss
 102  continue
      do 112 i=1,jj/2-1
              k=jj*0.5+i*1.0+1
              psd(i)=(data1(k+k-1)*data1(k+k-1)
```

```
     .                        +data1(k+k)*data1(k+k))/wss
 112    continue
c ********************************************************************
c
c  Smooth computed fft spectrum.
c
        if (idm(2).eq.1) then
         psdtemp=psd(1)
         do 820 i=1,jj-1
                psdavg=0.25*psdtemp+0.5*psd(i)+0.25*psd(i+1)
                psdtemp=psd(i)
                psd(i)=psdavg
 820     continue
         psd(jj)=0.25*psdtemp+0.75*psd(jj)
        endif
        do 850 i=1,jj
                if (psd(i).lt.1e-14) then
                  psddb(i)=-140.0
                else
                  psddb(i)=10*log10(psd(i))
                endif
 850    continue
c ********************************************************************
c
c Scaling for plotting
c Delta time = dt
c Delta freq = 1/(jj*dt)
c
        df=1/(jj*dt)
        do 142 i=1,nt
                xe(i)=(i*1.0-1)*dt
 142    continue
        do 145 i=1,jj
                xe1(i)=(j1+i-1)*dt
 145    continue
        do 111 i=1,jj
                xpsd(i)=(-jj*0.5+i)*df
 111    continue
c
c Plotting
c
        call linplot(xe,edata,nt,'Read-in electric field$',1)
        call linplot(xe1,efft,jj,'Selected electric field for FFT$',1)
        call linplot(xpsd,psd,jj,'PSD$',2)
        call dbplot(xpsd,psddb,jj,'PSD in dB$',2)
c ********************************************************************
c
c Magnifying plots for psd.
c
        fmax=0.5/dt
        write(*,*)'Max. plotting frequency = ', fmax
        write(*,*)
     .'(1) Magnify the whole spectrum by a magnifying ratio.'
        write(*,*)
     .'(2) Magnify part of the spectrum centered at an input freq.'
        write(*,*)'(3) Magnify some ranges of the spectrum.'
        write(*,*)'Select 1,2 or 3 :'
        read(*,*) ms
        if ((ms.lt.1).or.(ms.gt.3)) then
                write(*,*)'No such a selection!'
                goto 650
        endif
```

```
      go to (710,720,730), ms
c
c (1) Magnify the whole spectrum by a magnifying ratio.
c    (mr=magnifying ratio; npl=number of points for magnifying plotting)
c
 710   write(*,*)'What is the magnifying ratio(>1) for your psd plot?'
       read(*,*) mr
       mr=mr*2
       if (mr.le.1.0) goto 650
       if (mr.gt.jj/4.0) then
               write(*,*) 'Too high magnifying ratio!'
               goto 650
       endif
c
       ir=int(mr+0.5)
       npl=int((jj*1.0)/(ir*1.0))
       do 601 i=1,ir-1
       do 602 j=1,npl
               i1=(i-1)*npl+j
               xp(j)=xpsd(i1)
               mpsd(j)=psd(i1)
               mpsddb(j)=psddb(i1)
 602   continue
       call linplot(xp,mpsd,npl,'Magnified psd$',2)
       call dbplot(xp,mpsddb,npl,'Magnified psd in dB$',2)
 601   continue
       i2=jj-(ir-1)*npl
       if (i2.le.2) goto 650
       do 603 j=1,i2
               i1=(ir-1)*npl+j
               xp(j)=xpsd(i1)
               mpsd(j)=psd(i1)
               mpsddb(j)=psddb(i1)
 603   continue
       call linplot(xp,mpsd,npl,'Magnified psd$',2)
       call dbplot(xp,mpsddb,npl,'Magnified psd in dB$',2)
       goto 650
c
c (2) Magnify part of the spectrum centered at an input frequency.
c    (fc=center frequency; deltaf=frequency range for each plot;
c     np2=no. of deltaf you want on each side, including the plot
c        at center frequency.)
c
 720   write(*,*)'Input  fc, deltaf and # of deltaf (all > 0):'
       read(*,*) fc,deltaf,np2
       if ((fc.lt.0.0).or.(fc.gt.fmax)) then
               write(*,*)'Wrong fc!'
               goto 650
       endif
       if ((deltaf.le.0.0).or.((f+deltaf/2).ge.fmax)) then
               write(*,*)'Deltaf out of range!'
               goto 650
       endif
       if ((np2.lt.1).or.((fc+0.5*deltaf+np2*deltaf).ge.fmax)) then
               write(*,*)'Np2 out of range!'
               goto 650
       endif
c
       if1=int((fc-0.5*deltaf)/df+0.5)
       if2=int((fc+0.5*deltaf)/df+0.5)
       ideltaf=if2-if1
c
```

```
c  Left hand magnification
c
       do 721 i=2,np2
               k1=-(np2-i+1)*ideltaf+if1
               if ((k1+ideltaf).le.0) goto 721
       do 722 j=1,ideltaf
               k=k1+j-1+0.5*jj
               xp(j)=xpsd(k)
               mpsd(j)=psd(k)
               mpsddb(j)=psddb(k)
 722   continue
       call linplot(xp,mpsd,ideltaf,'Magnified psd$',2)
       call dbplot(xp,mpsddb,ideltaf,'Magnified psd in dB$',2)
 721   continue
c
c  Magnified plot at center frequency
c
       k1=if1
       do 723 j=1,ideltaf
               k=k1+j-1+0.5*jj
               xp(j)=xpsd(k)
               mpsd(j)=psd(k)
               mpsddb(j)=psddb(k)
 723   continue
       call linplot(xp,mpsd,ideltaf,'Magnified psd$',2)
       call dbplot(xp,mpsddb,ideltaf,'Magnified psd in dB$',2)
c
       fc1=xp(1)+int(0.5*ideltaf-0.5)*df
       do 726 j=1,ideltaf
               xp(j)=(j-1)*df-int(0.5*ideltaf-0.5)*df
 726   continue
       write(mlabel1,727) fc1
 727   format('Magnified psd (fc=',f5.3,')$')
       write(mlabel2,728) fc1
 728   format('Magnified psd in dB (fc=',f5.3,')$')
       call linplot(xp,mpsd,ideltaf,mlabel1,2)
       call dbplot(xp,mpsddb,ideltaf,mlabel2,2)
c
c  Right hand magnification
c
       do 724 i=2,np2
               k1=(i-1)*ideltaf+if1
               if (2*(k1+ideltaf).ge.jj) goto 724
       do 725 j=1,ideltaf
               k=k1+j-1+0.5*jj
               xp(j)=xpsd(k)
               mpsd(j)=psd(k)
               mpsddb(j)=psddb(k)
 725   continue
       call linplot(xp,mpsd,ideltaf,'Magnified psd$',2)
       call dbplot(xp,mpsddb,ideltaf,'Magnified psd in dB$',2)
 724   continue
       goto 650
c
c (3) Magnify some range of the spectrum.
c     (f1=Start frequency; f2=Stop frequency;
c      f1=0 and f2=0 means to end magnifying.)
c

 730   write(*,*)'Enter start and stop frequencies :'
       write(*,*)'(You may repeat as many times as you want;'
       write(*,*)' enter 0,0 to terminate magnifying section.)'
```

```
      fmin=-(jj/2-1)*df
      do 731 i=1,50
              read(*,*) f1,f2
              if ((f1.eq.0.0).and.(f2.eq.0.0)) goto 650
              if ((f1.lt.fmin).or.(f1.gt.fmax)) then
               write(*,*)'Start frequency out of range!'
               goto 731
              endif
              if ((f2.lt.fmin).or.(f2.gt.fmax)) then
               write(*,*)'Stop frequency out of range!'
               goto 731
              endif
              if (f2.lt.f1) then
               ftemp=f1
               f1=f2
               f2=ftemp
              endif
              if1=int(f1/df+0.5)
              if2=int(f2/df+0.5)
              ideltaf=if2-if1
      do 732 j=1,ideltaf
              k=if1+j-1+0.5*jj
              xp(j)=xpsd(k)
              mpsd(j)=psd(k)
              mpsddb(j)=psddb(k)
 732  continue
      call linplot(xp,mpsd,ideltaf,'Magnified psd$',2)
      call dbplot(xp,mpsddb,ideltaf,'Magnified psd in dB$',2)
 731  continue
c
c  End of magnification section.
c
 650  continue
      call CLSGKS
c ******************************************************************
c
c Write psddata1.dat
c
c       inquire(FILE='psddata1.dat', EXIST=data_ex)
c       if (data_ex.eqv..false.) go to 115
c              open(UNIT=4, FILE='psddata1.dat', STATUS='OLD')
c              close(UNIT=4, STATUS='DELETE')
c 115        continue
c       open(UNIT=6, FILE='psddata1.dat', STATUS='NEW')
c       do 118 i=1,jj
c              write(6,*) xpsd(i),psd(i),psddb(i)
c 118      continue
c       close(6)
c
      end
c ******************************************************************
c
c FFT subroutine (From Numerical Recipes, Chapter 12.)
c
      subroutine four1(data,nn,isign)
      common pi
      real*8 wr,wi,wpr,wpi,wtemp,theta
      dimension data(2*nn)
      n=2*nn
      k=1
      do 11 i=1,n,2
              if (k.gt.i) then
```

```
                        tempr=data(k)
                        tempi=data(k+1)
                        data(k)=data(i)
                        data(k+1)=data(i+1)
                        data(i)=tempr
                        data(i+1)=tempi
                endif
                m=n/2
 1              if ((m.ge.2).and.(k.gt.m)) then
                        k=k-m
                        m=m/2
                go to 1
                endif
                k=k+m
 11     continue
        mmax=2
 2      if (n.gt.mmax) then
                istep=2*mmax
                theta=2*pi/(isign*mmax)
                wpr=-2.d0*dsin(0.5d0*theta)**2
                wpi=dsin(theta)
                wr=1.d0
                wi=0.d0
                do 13 m=1,mmax,2
                        do 12 i=m,n,istep
                                k=i+mmax
                                tempr=sngl(wr)*data(k)-sngl(wi)*data(k+1)
                                tempi=sngl(wr)*data(k+1)+sngl(wi)*data(k)
                                data(k)=data(i)-tempr
                                data(k+1)=data(i+1)-tempi
                                data(i)=data(i)+tempr
                                data(i+1)=data(i+1)+tempi
12                      continue
                        wtemp=wr
                        wr=wr*wpr-wi*wpi+wr
                        wi=wi*wpr+wtemp*wpi+wi
13              continue
                mmax=istep
        go to 2
        endif
        return
        end
c ********************************************************************
c
c Plotting subroutine for psd in dB.
c  (Limited to 2000 points.)
c
        subroutine dbplot(xdata,ydata,nn,title,ilabel)
        common/chkrng/irs,dbmin,dbmax
        real xdata(65536),ydata(65536),xdata1(2050),ydata1(2050)
        integer nn,ntotal,nn1,k,k1,ilabel
        character*32 title,xlabel
c
        if (ilabel.eq.0) return
        xlabel=' $'
        if (irs.eq.1) then
          call agsetf('Y/MINIMUM.',dbmin)
          call agsetf('Y/MAXIMUM.',dbmax)
        endif
        ntotal=nn
        if (nn.gt.2000) then
          ntotal=0
```

```
          k=int(nn*1.0/2000)
          nn1=int((k*1.0+1.0)*(nn*1.0-1998.0*k)-k*1.0)
          k1=int((nn1*1.0-1.0)/(k*1.0+1.0)+1.0)
          nn1=int((k1*1.0-1.0)*(k*1.0+1.0)+1.0)
          do 100 i=1,nn1,(k+1)
            ntotal=ntotal+1
            xdata1(ntotal)=xdata(i)
            ydata1(ntotal)=ydata(i)
 100      continue
          do 120 i=(nn1+k),nn,k
            ntotal=ntotal+1
            xdata1(ntotal)=xdata(i)
            ydata1(ntotal)=ydata(i)
 120      continue
        else
          do 140 i=1,ntotal
            xdata1(i)=xdata(i)
            ydata1(i)=ydata(i)
 140      continue
        endif
        if (ilabel.eq.1) xlabel='Time$'
        if (ilabel.eq.2) xlabel='Frequency$'
        call agsetc('LABEL/NAME.','L')
        call agseti('LINE/NUMBER.',100)
        call agsetc('LINE/TEXT.',' $')
        call agsetc('LABEL/NAME.','B')
        call agseti('LINE/NUMBER.',-100)
        call agsetc('LINE/TEXT.',xlabel)
        call EZXY(xdata1,ydata1,ntotal,title)
        call agsetf('Y/MINIMUM.',1.e36)
        call agsetf('Y/MAXIMUM.',1.e36)
        return
        end
c *********************************************************************
c
c Plotting subroutine for psd in linear scale.
c  (Limited to 2000 points.)
c
        subroutine linplot(xdata,ydata,nn,title,ilabel)
        real xdata(65536),ydata(65536),xdata1(2050),ydata1(2050)
        integer nn,ntotal,nn1,k,k1,ilabel
        character*32 title,xlabel
c
        if (ilabel.eq.0) return
        xlabel=' $'
        ntotal=nn
        if (nn.gt.2000) then
          ntotal=0
          k=int(nn*1.0/2000)
          nn1=int((k*1.0+1.0)*(nn*1.0-1998.0*k)-k*1.0)
          k1=int((nn1*1.0-1.0)/(k*1.0+1.0)+1.0)
          nn1=int((k1*1.0-1.0)*(k*1.0+1.0)+1.0)
          do 100 i=1,nn1,(k+1)
            ntotal=ntotal+1
            xdata1(ntotal)=xdata(i)
            ydata1(ntotal)=ydata(i)
 100      continue
          do 120 i=(nn1+k),nn,k
            ntotal=ntotal+1
            xdata1(ntotal)=xdata(i)
            ydata1(ntotal)=ydata(i)
 120      continue
```

```
      else
        do 130 i=1,ntotal
          xdata1(i)=xdata(i)
          ydata1(i)=ydata(i)
130     continue
      endif
      if (ilabel.eq.1) xlabel='Time$'
      if (ilabel.eq.2) xlabel='Frequency$'
      call agsetc('LABEL/NAME.','L')
      call agseti('LINE/NUMBER.',100)
      call agsetc('LINE/TEXT.',' $')
      call agsetc('LABEL/NAME.','B')
      call agseti('LINE/NUMBER.',-100)
      call agsetc('LINE/TEXT.',xlabel)
      call EZXY(xdata1,ydata1,ntotal,title)
      return
      end
```

(C) Input/output file formats and examples

The ES1G program requires only one input file (ES1G.DAT) to set the simulation parameters. The ES1G.DAT file uses text format with the following input sequence of parameters.

$lll \;\; nsp \;\; dt \;\; nt \;\; ng \;\; e_o \;\; \omega_o \;\; iplot \;\; \theta_B \;\; mplot(7)$

$N \;\; \omega_{po} \;\; \Omega \;\; qm \;\; v_t \;\; v_o \;\; mode \;\; x_1 \;\; v_1 \;\; x_{min} \;\; x_{max} \;\; dfn$

$\vdots$

$ngavg \;\; ngpsd \;\; itavg \;\; ixavg$

$ip(9)$

$iani \;\; iheight \;\; iwidth \;\; ipicture$

$itrace \;\; xo_{min} \;\; xo_{max} \;\; npt$

The number of parameter sets of each species corresponds to the number of species (nsp). The parameters, *iplot* and *mplot*(7), specify the time-step interval for diagnostic plots and Fourier mode numbers for history plots, respectively. The parameter, v_o, is the initial drift velocity of the species, which is not implemented in the ES1G program. The two parameters, *itavg* and *ixavg*, are used for time and spatial averaging of density plots. The array, *ip*(9), specifies which phase space and velocity distribution plots of all species are required. The density animation and particle trace plots can be activated by setting *iani* and *itrace* to 1. The three parameters, *iheight*, *iwidth* and *ipicture*, are associated with the setup of the density animation, where *iheight* and *iwidth* respectively specify the number of vertical and horizontal pixels, and *ipicture* is the number of frames to show in the animation. More detail about other input parameters can be found in the ES1G program listing and [1] Birdsall. As an example, the input file of the reference case ($\omega_o = 0.88$) is depicted below.

```
75 2 0.2 20000 1024 0.15 0.88 154 31 1000 0. 0 0 0 0 0 0 0 0
40000 1.172 -0.3 -1. 1.0 0. 1 0. 0. 0.05 0.95 1.314
40000 0.167429 0.00612245 0.020408 0.0714286 0. 1 0. 0. 0.05 0.95 1.314
7 154 9 3
1 0 0 0 0 0 0 0 0
1 64 256 220
1 0.1 0.9 25
```

Note that the numerical accuracy of the plasma frequency and q/m ratio of each species is important for the charge neutrality of the whole plasma. At least five significant digits should be used. Figure C-1(c) shows the effect of the charge difference between the two species. This in turn causes a finite electric field other than the pump region in Figure C-1(e).

After invoking the ES1G program, three files are generated. The ES1G.DAT file contains the electric field data at each time step for the calculation of power spectrum. Its file format is,

dt nt ngpsd
First electric field data
⋮
Last electric field data

The data type of the second and third numbers is integer, while the data type of the time step and the electric field data is single-precision real. Note that *ngpsd* and *ngavg* specify the observation position and the number of adjacent grid points to be averaged for the output electric field. The next data file produced by the ES1G program is the file ES1G.IMAGE which is converted to another file ES1G.HDF using the HDFCONV program. It is then used for density animation. The purpose of such an animation is to investigate the development and propagation of plasma tubulence generated by the nonlinear interaction between the pump and the plasma. The format of the file ES1G.IMAGE is,

iy ix it iblack iwhite

$$\left.\begin{array}{l}\vdots\\ \text{(Normalized density for the first species)}\\ \vdots\end{array}\right\}(ix-1)/2\ \ \text{data}$$

1000
⋮
(Normalized density for the second species)
⋮

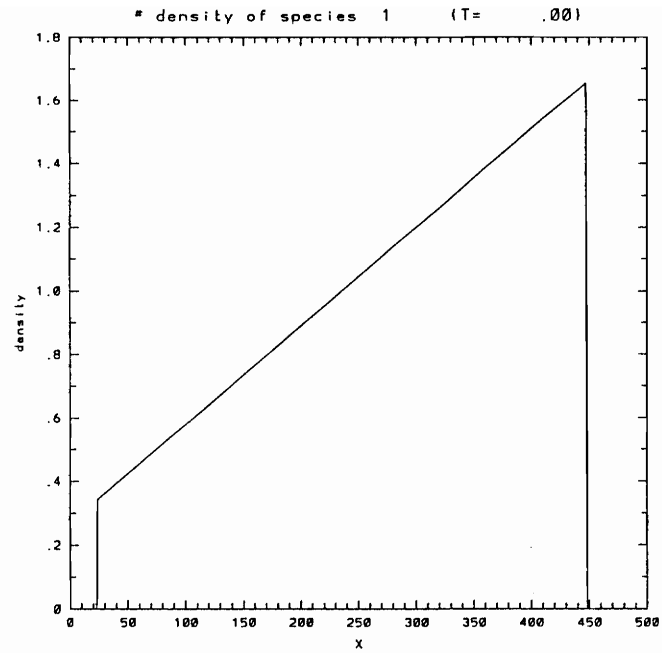
⋮

All data contained in the file ES1G.IMAGE are integers. The data of the normalized density range from 1 to the number of vertical pixels ($iy = iheight$). The two density profiles share a frame horizontally, with a blank vertical line in between. The total number of horizontal pixels is $ix = iwidth$. The number "1000" which appears after the last density data of the first species for each frame, indicates a separation of the two sets of data. At each horizontal pixel, it corresponds to a density value. Then, the HDFCONV program fills the density profiles with the predefined the black and white levels by the two numbers, *iblack* and *iwhite*. The total number of pictures is specified by $it = ipicture$. With the sample settings in the above ES1G.INPUT file, the frame size is 64×257[1] and the number of pictures is 220, which is barely enough to obtain a coherent movie of plasma tubulence.

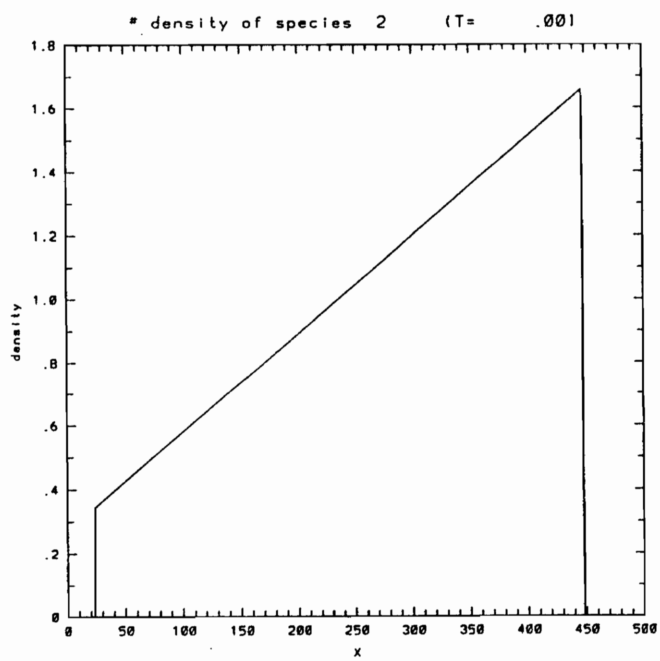
The third output file from the ES1G program is the GMETA file, which contains all the diagnostic and history plots. Subroutines from NCAR Graphics are used in the ES1G program to generate these plots. They can be directly viewed using a graphic terminal, as well as they can be plotted to a printer. Figure C-1 and C-2 are some snapshots of the normalized particle densities, charge density, potential, electric field, and phase spaces of the reference case at $t = 0$ and $t = 1400$. For the example settings, the ES1G program generates these diagnostic plots every 1000 time steps or 200 time units. Figure C-3 depicts the corresponding history plots.

Finally, as an example illustration of how to use the PSD program, Figure 23(a) and (b) are generated by the following selections of the options provided by the PSD program. After invoking the PSD program, the data manipulation option *(2) Smooth computed fft spectrum* is chosen. Then, the PSD program will display a text plot of the electric field time series and asks for inputting the start time and end time of the data segment for FFT. To generate our desired spectrum, one may enter *600,3800*. The PSD program goes on to the data windowing and dB range selections. Choose *(1) Square* and then *(1) Auto range*. The last selection needed to accomplish the run is the styles of the magnified plots. Since we want a spectral plot centered at the pump frequency, *(2) Magnify part of the spectrum centered at an input freq.* should be selected. After that, enter the center frequency (f_c), the total frequency deviation (*Deltaf*) and the number of Δf expansions around the center frequency as *0.14,0.06,3*. After the PSD program is terminated, its output plots are contained in the GMETA file.

[1] The ES1G program will automatically add one to the parameter *iwidth* to make it odd.

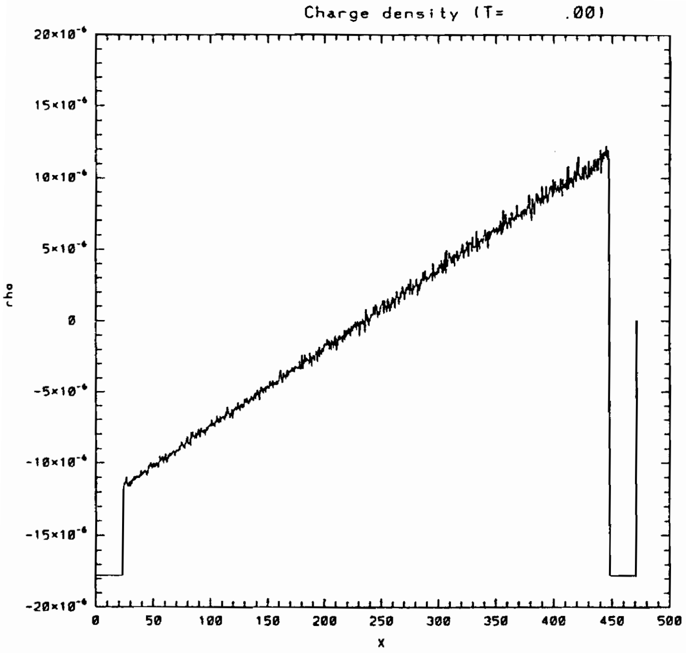


(a) Electron density

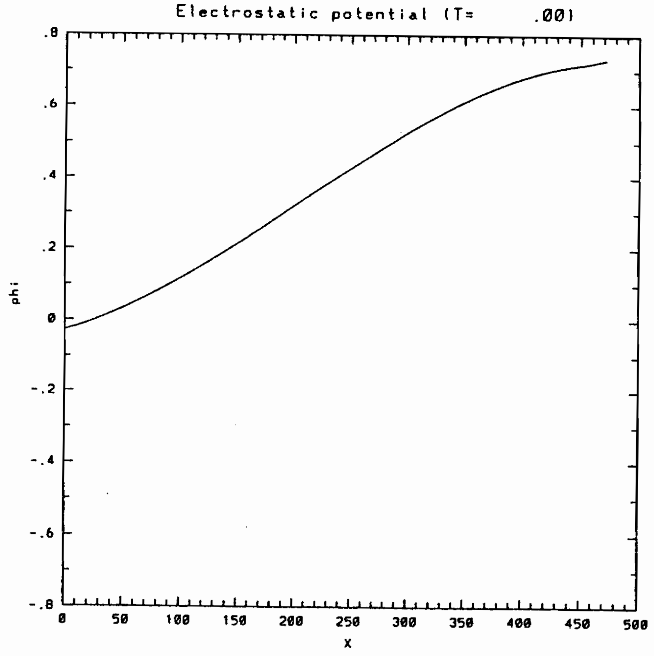


(b) Ion density

Figure C-1: Diagnostic plots of the reference case at $t = 0$. (Initial loading)

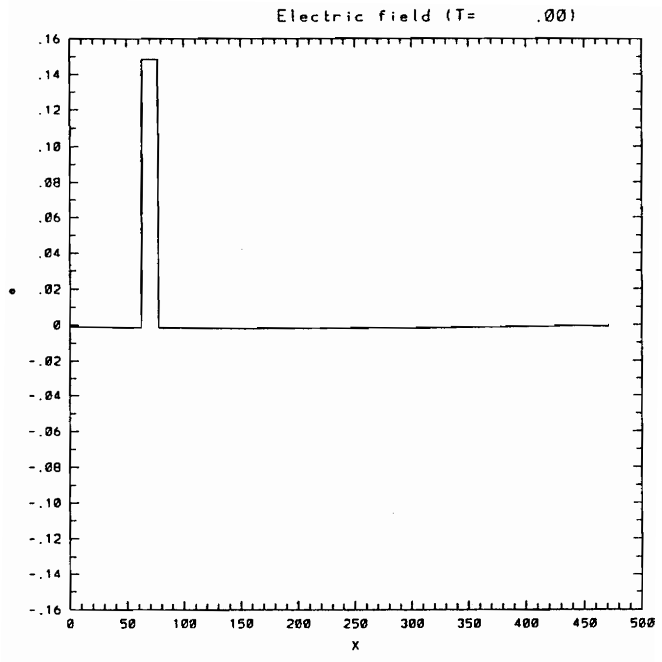


(c) Charge density

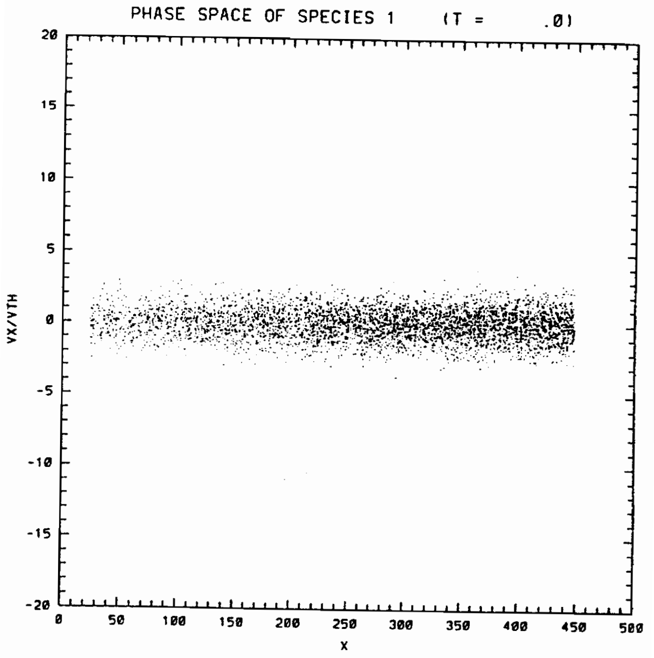


(d) Potential

Figure C-1 (continued).

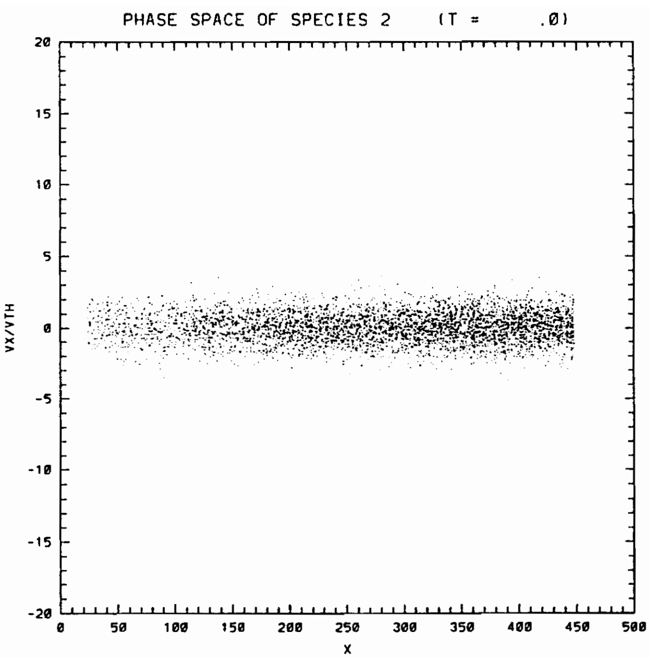


(e) Electric field



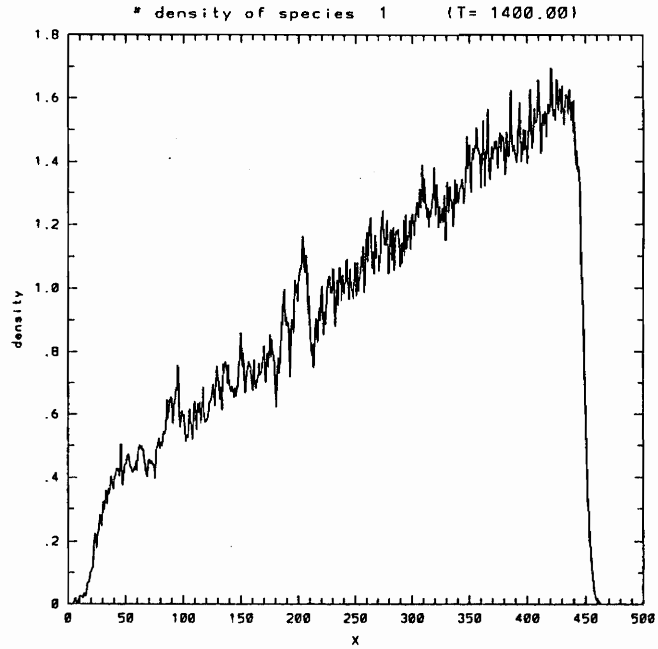
(f) Electron phase space

Figure C-1 (continued).

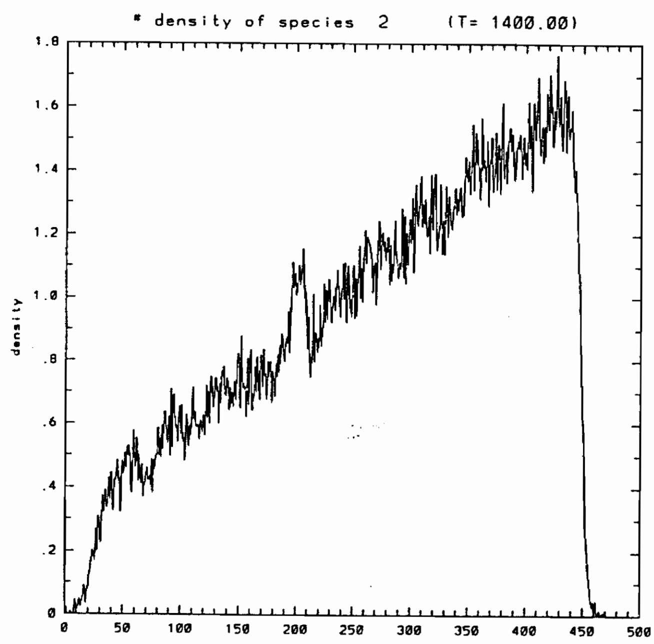


(g) Ion phase space

Figure C-1 (continued).

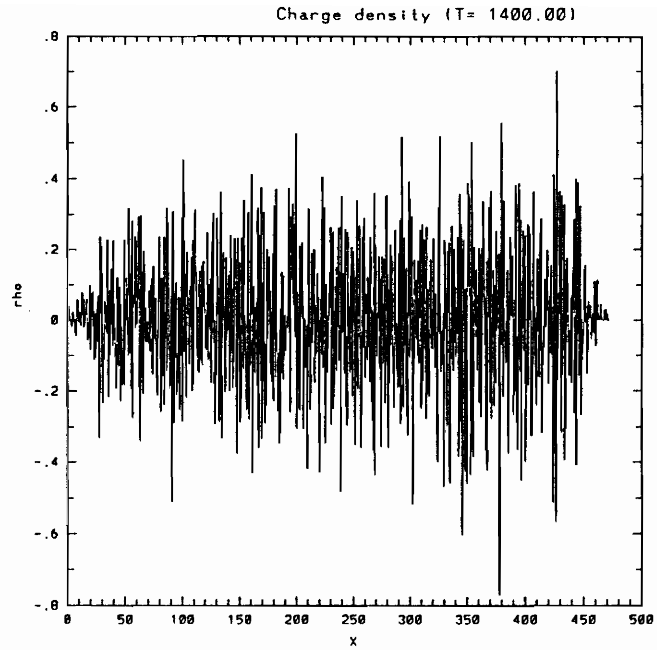


(a) Electron density

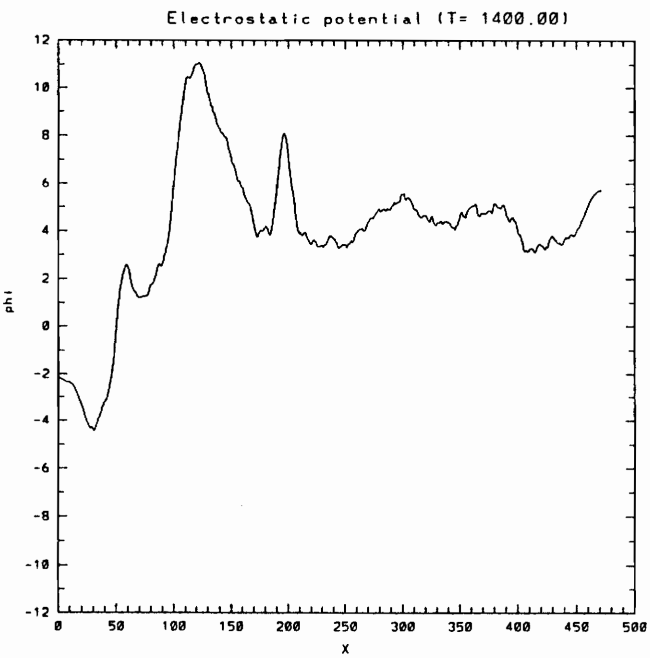


(b) Ion density

Figure C-2: Diagnostic plots at $t = 1400$.

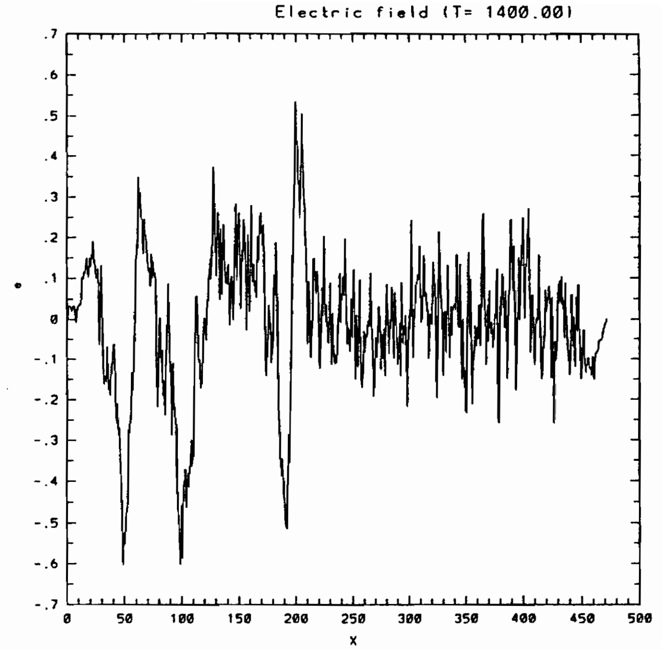


(c) Charge density

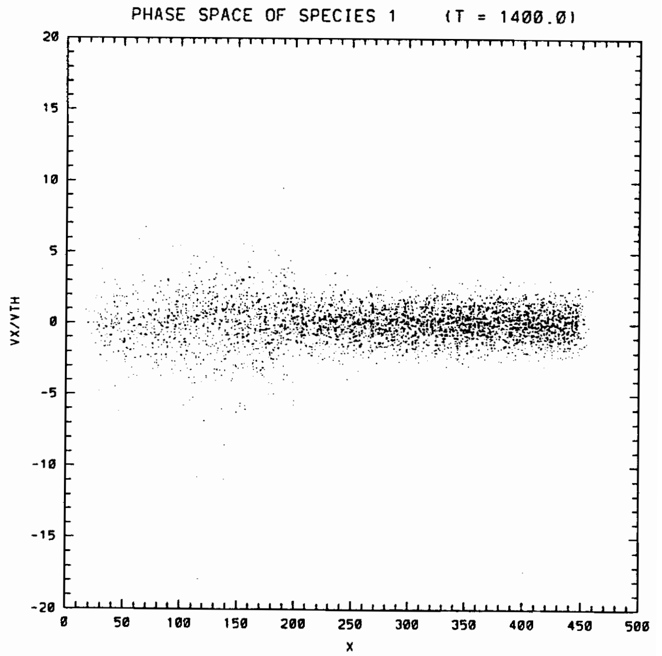


(d) Potential

Figure C-2 (continued).

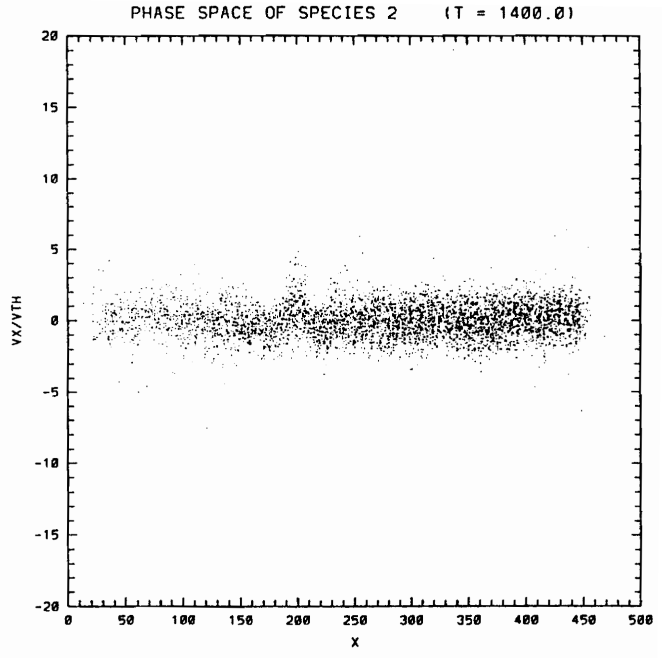


(e) Electric field



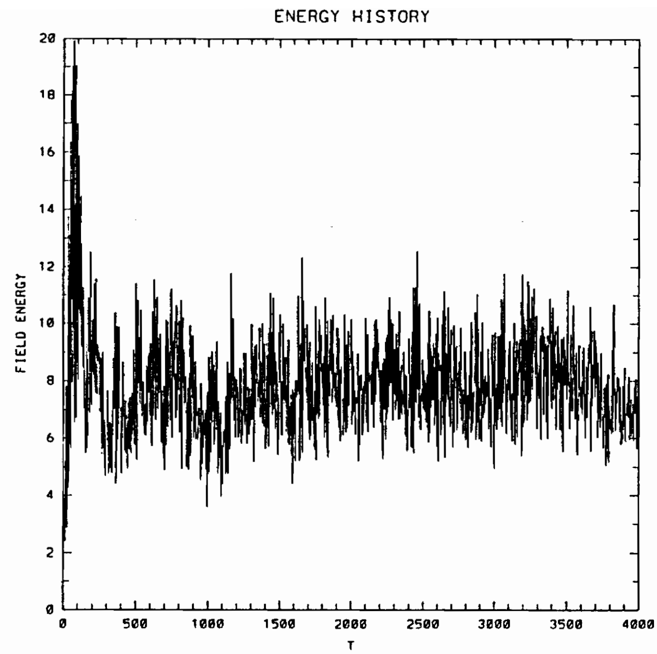
(f) Electron phase space

Figure C-2 (continued).

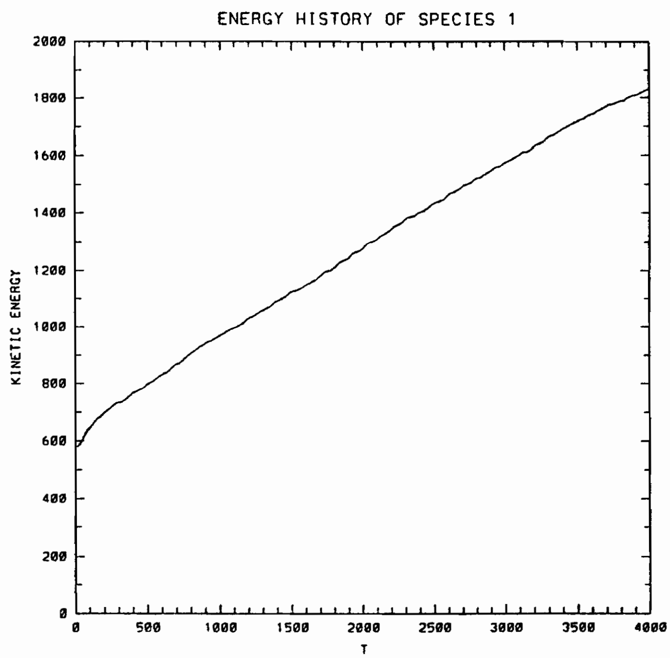


(g) Ion phase space

Figure C-2 (continued).

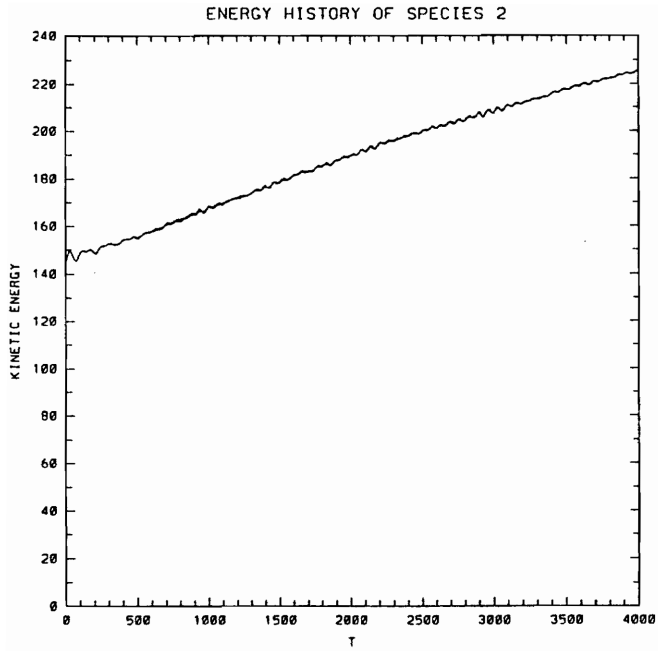


(a) Electric field energy

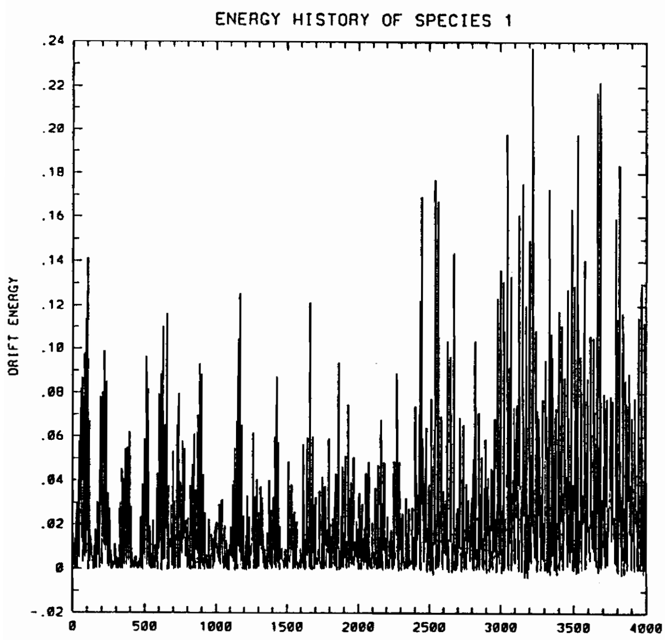


(b) Kinetic energy of electrons

Figure C-3: History plots.

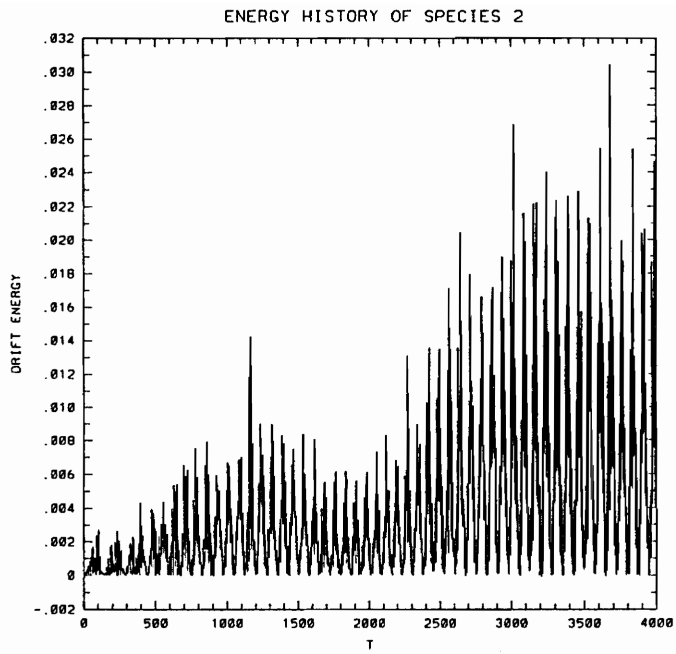


(c) Kinetic energy of ions

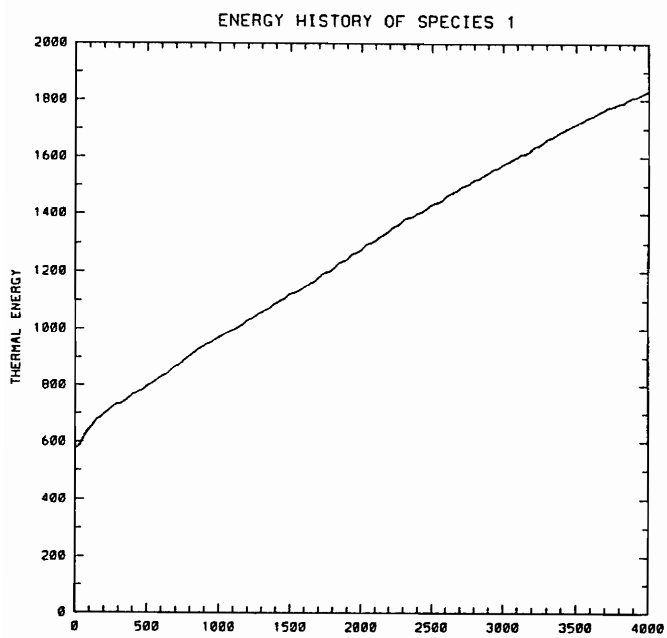


(d) Drift energy of electrons

Figure C-3 (continued).

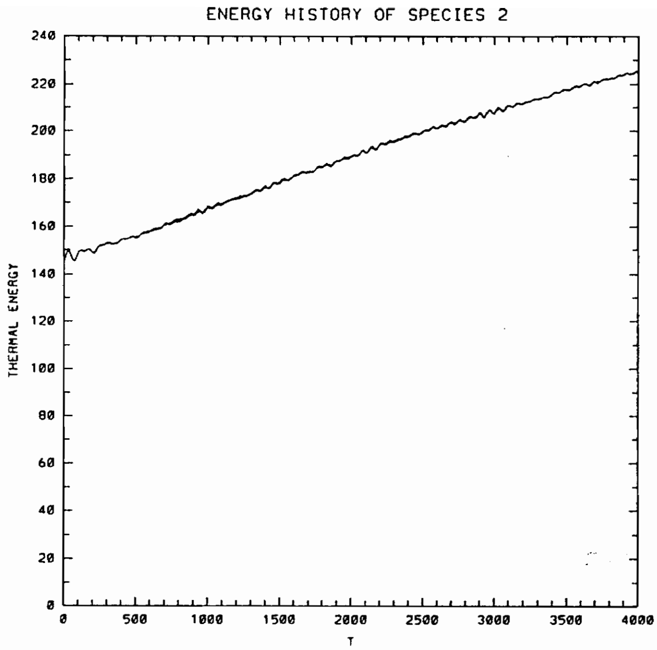


(e) Drift energy of ions

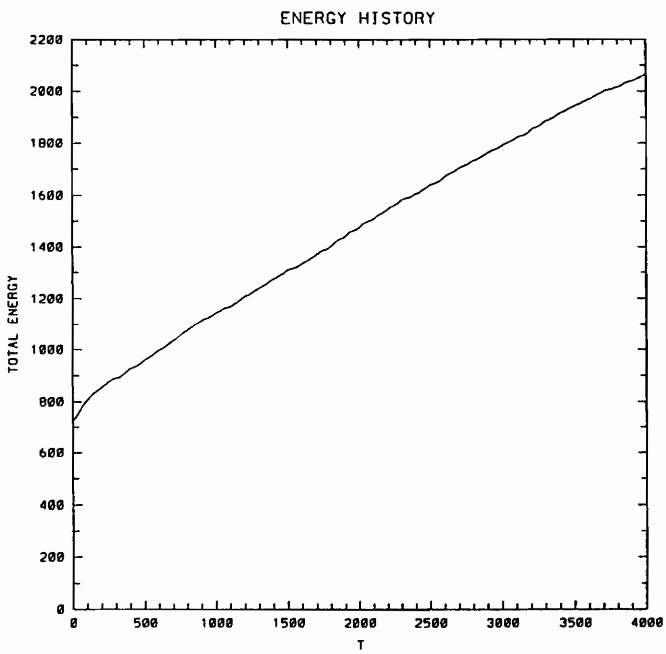


(f) Thermal energy of electrons

Figure C-3 (continued).

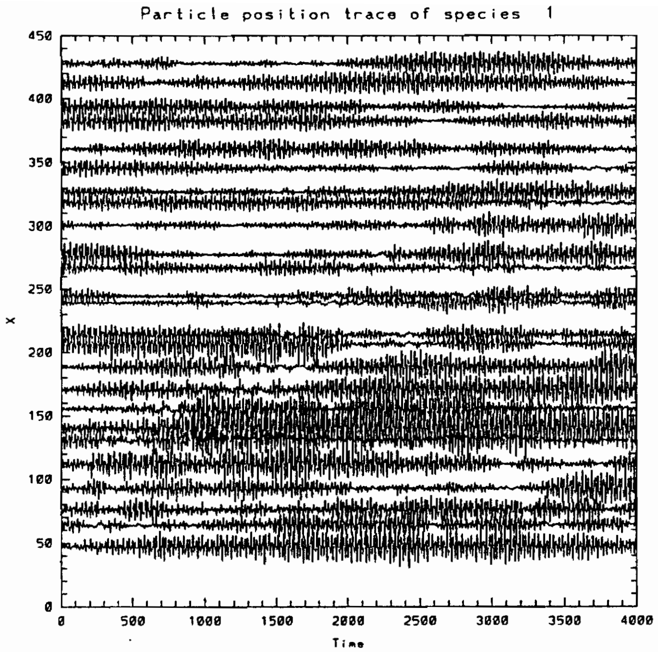


(g) Thermal energy of ions

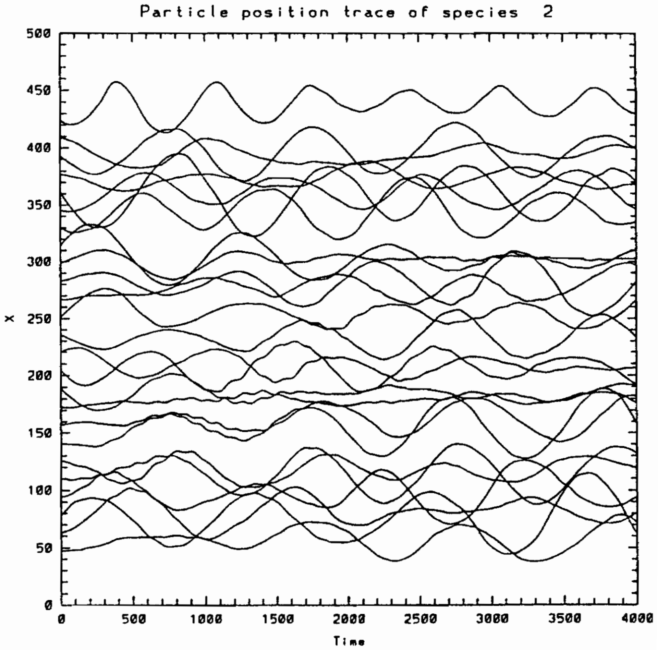


(h) Total energy

Figure C-3 (continued).



(i) Electron trajectories



(j) Ion trajectories

Figure C-3 (continued).

(D) List of simulated cases

More than 70 cases have been simulated in this research work. They are summarized in this appendix for future reference, including the cases depicted in Chapter 5. Most of the simulated results (including the GMETA files containing disgnostic plots, ES1G.DAT files, and ES1G.IMAGE files) are stored in a tape and some of important plots are kept in files[1]. The date of simulation run is embedded in their reference name. The simulation parameters are assumed to follow the settings according to Table 3, unless specified otherwise.

Case	ω_o	lll	ω_{peo}	$ecenter$	Remarks
es1gf0515	0.0	75	1.172	154	eo=0.0 (Noise case)
es1ge0519	0.58	175.6	0.766	107	
es1gm0521	0.58	286.2	0.6	287	
es1gc0516	0.84	75	1.172	119	
es1gb0516	0.86	75	1.172	136	
es1gi0520	0.87	75	1.172	145	
es1gh0520	0.875	75	1.172	150	
es1ga0511	0.88	100	1.172	154	espan=41
es1gb0511	0.88	75	1.172	154	espan=41
es1gc0511	0.88	50	1.172	154	espan=41
es1gd0512	0.88	65	1.172	154	espan=41
es1ge0512	0.88	75	1.172	154	espan=81
es1gf0512	0.88	75	1.172	154	espan=101
es1gg0512	0.88	75	1.172	154	espan=61
es1gh0512	0.88	75	1.172	154	espan=21
es1gj0514	0.88	75	1.172	154	ngavg=9
es1gk0514	0.88	75	1.172	154	ngavg=5
es1gl0514	0.88	75	1.172	154	ngavg=3
es1ga0515	0.88	75	1.172	154	eo=0.2
es1gc0515	0.88	75	1.172	154	eo=0.05
es1gd0515	0.88	75	1.172	154	eo=0.3
es1ge0515	0.88	75	1.172	154	eo=0.4
es1ga0516	0.88	75	1.172	154	Reference case
es1ga0522	0.88	106.8	0.9	393	
es1ge0522	0.88	75	1.172	157	K=100
es1gg0524	0.88	75	1.172	159	K=225, dt=0.18, nt=25000
es1ga0606	0.88	75	1.172	308	ng=2048
es1gc0606	0.88	75	1.172	154	ngpsd=104
es1gd0606	0.88	75	1.172	154	ngpsd=204
es1ge0607	0.88	75	1.172	154	ngpsd=146
es1gf0607	0.88	75	1.172	154	ngpsd=124
es1gd0607	0.88	75	1.172	154	ngpsd=64

[1]Both of them are available from Dr. W.A. Scales. Interested readers may contact him at Room 615, Whittemore Hall, Department of Electrical Engineering, Virginia Tech, Blacksburg, VA 24061. Also, for more details about the input parameters, please contact Dr. Scales for the simulation log sheets.

es1ga0608	0.88	106.8	0.9	393	ngpsd=343
es1gb0608	0.88	106.8	0.9	393	ngpsd=443
es1gc0608	0.88	106.8	0.9	393	ngpsd=293
es1gd0609	0.88	106.8	0.9	199	dfn=0.5
es1ga0610	0.88	75	1.172	154	$v_{te} = 0.5$
es1gb0610	0.88	75	1.172	154	$v_{ti} = 0.0357143$
es1gc0610	0.88	75	1.172	154	$v_{ti} = 0.1428572$
es1ga0617	0.88	75	1.172	154	$v_{te} = 2$
es1gg0520	0.885	75	1.172	158	
es1gd0516	0.89	75	1.172	163	
es1ge0516	0.9	75	1.172	172	Third harmonic case
es1gb0606	0.9	75	1.172	172	$\theta_B = 8.19$
es1gd0522	0.905	75	1.172	176	
es1gf0517	0.91	75	1.172	181	
es1gh0524	0.91	75	1.172	186	K=225, dt=0.18, nt=25000
es1gd0611	0.91	75	1.172	181	$v_{te} = 0.5$
es1ge0611	0.91	75	1.172	181	$v_{te} = 2$
es1gf0611	0.91	75	1.172	181	$v_{ti} = 0.0357143$
es1gg0517	0.92	75	1.172	190	
es1gf0522	0.92	75	1.172	193	K=100
es1gh0517	0.94	75	1.172	209	
es1gi0517	0.96	75	1.172	228	
es1gj0517	1.125	75	1.172	400	Corresponds to 5.1MHz
es1gm0518	1.14	75	1.172	417	
es1gl0518	1.16	75	1.172	440	
es1gk0518	1.18	75	1.172	463	
es1gb0522	1.18	33.8	1.6	161	
es1gi0524	1.18	75	1.172	473	K=225, dt=0.18, nt=25000
es1ge0609	1.18	100	1.0	707	
es1gj0520	1.185	75	1.172	469	
es1gn0518	1.19	75	1.172	475	
es1gk0521	1.195	75	1.172	481	
es1go0518	1.2	75	1.172	487	Fourth harmonic case
es1gl0520	1.205	75	1.172	493	
es1gc0518	1.205	33.8	1.6	177	
es1ga0519	1.21	75	1.172	499	
es1gb0519	1.22	75	1.172	511	
es1gf0609	1.22	100	1.0	773	
es1ga0611	1.22	75	1.172	511	ngpsd=471
es1gb0611	1.22	75	1.172	511	ngpsd=551
es1gc0519	1.24	75	1.172	536	
es1gd0519	1.26	75	1.172	561	
es1gf0519	1.48	51.1	1.42	573	

(E) UH program listing

The program UH is used to calculate the upper hybrid point (in terms of the grid position) and the lower hybrid frequency for a given pump frequency. Its main purpose is to facilitate the setting of *ecenter* for the SEE simulation using the program ES1G. Note that in the program UH, the variables, k_l and *ngo* refer to the normalized effective plasma length and the grid point at the center of the effective plasma, respectively. Here is the listing of the program UH.

```
c ****************************************
c   program UH.F
c
c   (Calculation of upper hybrid point
c    for a given pump frequency.)
c
c   BY  K.T. Cheng     (20 FEB 1993)
c ****************************************
c
      real K,w,wce,wci,wpe,a,b,wpi,wlh,f,flh
      real kl,ng,dfn,wpeo,ngo,ng2,twopi
      integer is
      twopi=6.283 185 307
c
c
      write(*,*)'(1) Use default values'
      write(*,*)'    (kl=0.9,dfn=1.314,ngo=512,ng=1024,wpeo=1.172)'
      write(*,*)'(2) Input your own parameters'
      write(*,*)'Select 1 or 2 :'
      read(*,*) is
      if ((is.ne.1).and.(is.ne.2)) stop
      go to (100,200), is
 100  kl=0.9
      dfn=1.314
      ngo=512
      ng=1024
      wpeo=1.172
      goto 300
 200  write(*,*)'Input kl,dfn,ngo,ng,wpeo :'
      read(*,*) kl,dfn,ngo,ng,wpeo
c
 300  continue
      write(*,*)'Input K,w,wce :'
      read(*,*) K,w,wce
      wci=-wce/K
      wpe=sqrt((w*w-wce*wce)/(1+1/K-(wce/w)*(wce/w)/K))
      wpi=wpe/sqrt(K)
      a=0.5*(wpe*wpe+wce*wce+wpi*wpi)
      b=sqrt((0.5*(wpe*wpe+wce*wce-wpi*wpi))**2+(wpe*wpi)**2)
      wlh=sqrt(a-b)
      f=w/twopi
      flh=wlh/twopi
      ng2=ngo+ng*kl*((wpe/wpeo)**2-1)/dfn
c
c
      write(*,*)'f=', f
      write(*,*)'wpe=',wpe
```

```
write(*,*)'wlh=',wlh
write(*,*)'flh=',flh
write(*,*)'ng2=',ng2
write(*,*)'wci=',wci
end
```

(F) HDFCONV program listing

It is used to convert a ES1G.IMAGE file to a ES1G.HDF file for density animation. Note that for a $64 \times 257 \times 220$ animation, the file size of ES1G.IMAGE and ES1G.HDF are typically 350 Kbytes and 9.8 Mbytes, respectively. For storage purpose, it is recommended to use the IMAGE files only. The listing of the program HDFCONV is below.

```
c ****************************************************
c program HDFCONV.F
c ****************************************************
c
c  By K.T. Cheng
c
c  For conversion of ES1G.IMAGE to ES1G.HDF for density
c  animation (xds).
c ****************************************************
      real image(256,512,128)
c
c
      open(5,file='es1g.image',status='old',err=888)
      read(5,*) iy,ix,it,iblack,iwhite
      call hdf(image,ix,iy,it,iblack,iwhite)
      close(5)

 888  stop
      end
c *******************************************************************
c
      subroutine hdf(image,ix,iy,it,iblack,iwhite)
      real image(ix,iy,it)
      integer DFSDsetdims,DFSDputdata,ret
      integer ix,iy,it,temp,shape(3)

      shape(1)=ix
      shape(2)=iy
      shape(3)=it
      ix1=0.5*(ix-1)
      do 100 k=1,it
      do 120 i=1,ix1
      read(5,*) iy1
      do 140 j=1,iy1
        image(i,j,k)=iwhite
 140  continue
      do 150 j=iy1+1,iy
        image(i,j,k)=iblack
 150  continue
 120  continue
c
      read(5,*) iy1
      if (iy1.ne.1000) then
        write(*,*) 'Error in ES1G.IMAGE file.'
        stop
      endif
      do 160 j=1,iy
        image(ix+1,j,it)=255
 160  continue
```

```
c
      do 220 i=ix1+2,ix
      read(5,*) iy1
      do 240 j=1,iy1
        image(i,j,k)=iwhite
 240  continue
      do 250 j=iy1+1,iy
        image(i,j,k)=iblack
 250  continue
 220  continue
 100  continue
c
c  Write image to HDF file.
c
      ret=DFSDsetdims(3,shape)
      ret=DFSDputdata('es1g.hdf',3,shape,image)
      if (ret.ne.0) then
        write(*,*)'Error writing HDF file.'
      endif
      return
      end
```

References

Books/dissertations :

[1] Birdsall C.K. and Langdon A.B., Plasma physics via computer simulation (McGraw-Hill 1985)

[2] Budden K.G., The propagation of radio waves (Cambridge University Press 1985)

[3] Chen F.F., Introduction to plasma physics and controlled fusion, 2nd edition (Plenum Press 1984)

[4] Davies K., Ionospheric radio (Peter Peregrinus 1990)

[5] Hargreaves J.K., The solar-terrestrial environment (Cambridge University Press 1992)

[6] Hines C.O. *et al* (edited), Physics of the Earth's upper atmosphere (Prentice Hall 1965)

[7] Hockey R.W. and Eastwood J.W., Computer simulation using particles (Adam Hilger 1988)

[8] Isham B.C., Chirped incoherent scatter radar observations of the HF-modified ionosphere (Dissertation, Cornell University 1991)

[9] Kelley M.C., The Earth's ionosphere : Plasma physics and electrodynamics (Academic Press 1989)

[10] Kruer W.L., The physics of laser plasma interactions (Addison-Wesley 1988)

[11] Leyser T.B., Stimulated electromagnetic emission in the ionosphere (Dissertation, Swedish Institute of Space Physics 1989)

[12] Nicholson D.R., Introduction to plasma theory (John Wiley & Sons 1983)

[13] Press W.H. *et al*, Numerical recipes : The art of scientific computing (FORTRAN), (Cambridge University Press 1989)

[14] Reitz J.R. and Milford F.J., Foundation of electromagnetic theory, 2nd edition (Addison-Wesley 1967)

[15] Schmidt G., Physics of high temperature plasmas : An introduction (Academic Press 1966)

[16] Stix T.X., Waves in plasmas (American Institute of Physics 1992)

[17] Swanson D.G., Plasma waves (Academic Press 1989)

Published papers :

[1] Antani S.N. *et al* (1991 Dec) "Direct conversion of ordinary mode into upper hybrid wave by density irregularities in the ionosphere", Geophysical Research Letters (Vol.18, No.12, p.2285-2288)

[2] Armstrong W.T. *et al* (1990 Nov) "Continuous measurement of stimulated electromagnetic emission spectra from HF excited ionospheric turbulence", Radio Science (Vol.25, No.6, p.1283-1289)

[3] Boiko G.N. (1985) "Dynamic characteristics of stimulated radio emission", Radiophysics (Vol.28, No.4, p.259-268)

[4] Carlson H.C. (1990) "High power HF modification : Geophysics", published in *Ionospheric modification and its potential to enhance or degrade the performance of military systems* IB, AGARD (p.IB-1 – IB-13)

[5] Close R.A. *et al* (1990 Nov) "Computer simulation of ionospheric radio frequency heating", Radio Science (Vol.25, No.6, p.1341-1349)

[6] Goodman S. (1991 Dec) "Stimulated electromagnetic emissions from magnetized and inhomogeneous plasma", Journal of Geophysical Research (Vol.96, No.A12, p.21291-21298)

[7] Huang J. *et al* (1993 May) "The effect of electron cyclotron harmonic resonance on the frequency spectrum of the HF-heater-induced stimulated electromagnetic emissions (SEEs)", Proceedings of the 7th International Ionospheric Effects Symposium (p.6B-3-1 to 6B-3-8)

[8] Jones T.B. (1990) "The physics of ground based heating", published in *Ionospheric modification and its potential to enhance or degrade the performance of military systems* IA, AGARD (p.IA-1 – IA-9)

[9] Leyser T.B. and Thide B. (1988 Aug) "Effect of pump-induced density depletions on the spectrum of stimulated electromagnetic emissions", Journal of Geophysical Research (Vol.93, No.A8, p.8681-8688)

[10] Leyser T.B. *et al* (1989 Sep) "Stimulated electromagnetic emission near electron cyclotron harmonics in the ionosphere", Physical Review Letters (Vol.63, No.11, p.1145-1147)

[11] Leyser T.B. *et al* (1990 Oct) "Dependence of stimulated electromagnetic emission on the ionosphere and the pump wave", Journal of Geophysical Research (Vol.95, No.A10, p.17233-17244)

[12] Leyser T.B. (1991 Mar) "Parametric interaction between upper hybrid and lower hybrid waves in heating experiments", Geophysical Research Letters (Vol.18, No.3, p.408-411)

[13] Leyser T.B. *et al* (1992 June) "Narrow cyclotron harmonic absorption resonances of stimulated electromagnetic emission in the ionosphere", Physical Review Letters (Vol.68, No.22, p.3299-3302)

[14] Lin A.T. *et al* (1982 Apr) "Plasma heating from upper-hybrid mode conversion in an inhomogeneous magnetic field", Physics of Fluids (Vol.25, No.4, p.646-651)

[15] Mjolhus E. and Fla T. (1984 June) "Direct access to plasma resonance in ionospheric radio experiments", Journal of Geophysical Research (Vol.89, No.A6, p.3921-3928)

[16] Mjolhus E. (1990 Nov) "On linear conversion in a magnetized plasma", Radio Science (Vol.25, No.6, p.1321-1339)

[17] Oya H. (1971 Dec) "Conversion of electrostatic plasma waves into electromagnetic waves : Numerical calculation of the dispersion relation for all wavelengths", Radio Science (Vol.6, No.12, p.1131-1141)

[18] Rao N.N. and Kaup D.J. (1990 Oct) "Upper hybrid mode conversion and resonance excitation of Berstein modes in ionospheric heating experiments", Journal of Geophysical Research (Vol.95, No.A10, p.17245-17252)

[19] Rao N.N. and Kaup D.J. (1992 May) "Excitation of electron cyclotron harmonic waves in ionospheric modification experiments", Journal of Geophysical Research (Vol.97, No.A5, p.6323-6341)

[20] Stubbe P. *et al* (1984 Sep) "Stimulated electromagnetic emission: A new technique to study the parametric decay instability in the ionosphere", Journal of Geophysical Research (Vol.89, No.A9, p.7523-7536)

[21] Stubbe P. and Kopka H. (1990 July) "Stimulated electromagnetic emission in a magnetized plasma : A new symmetric special feature", Physical Review Letters (Vol.65, No.2, p.183-186)

[22] Thide B. *et al* (1982 Nov) "Observations of stimulated scattering of a strong high-frequency radio wave in the ionosphere", Physical Review Letters (Vol.49, No.21, p.1561-1564)

[23] Thide B. and Hedberg A. (1989 May) "First observations of stimulated electromagnetic emission at Arecibo", Geophysical Research Letters (Vol.16, No.5, p.369-372)

[24] Wong A.Y. *et al* (1981 Nov) "Rapid conversion of electromagnetic waves in the ionosphere", Physical Review Letters (Vol.47, No.18, p.1340-1343)

[25] Wong A.Y. *et al* (1987 Mar) "Observation of ionospheric cavitons", Physical Review Letters (Vol.58, No.13, p.1375-1378)

Vita

Cheng Kim-tung (K.T. Cheng) was born in Hong Kong, March 30, 1967. He graduated from the Hong Kong Polytechnic in July 1990 and received his B.Eng. degree with the first class honor, in Electronic Engineering. In 1988 to 1989, he worked in Motorola Semiconductors Hong Kong Limited as a cooperative student. After his graduation, he was employed as a senior engineer in the Advanced Engineering Division, Astec International Limited before he came to the United States in 1991 for his master study.

In August 1991, Mr. Cheng first joined the Virginia Power Electronic Center and then in June 1992, the electromagnetic research group of Virginia Tech. He worked on the numerical simulation of stimulated electromagnetic emissions in the ionosphere as his master thesis topic. The areas of his research interests are broad, including high-frequency high-density switch-mode power conversions, nonlinear characterization and modelling of magnetic materials, winding design of inductors and transformers, analog and power electronic circuit designs, plasma simulation, ionospheric propagation, and the philosphy behind Newton's Laws, Maxwell's Equations, quantum mechanics and relativity. He plans to devote his lifetime to search for the meaning of life.

K.T. Cheng

K.T.Cheng