Frequency Analysis of Low-Frequency Field Fluctuations Detected by the Cluster Satellites

Yiwen Xu

Master of Science Thesis
Stockholm, Sweden 2007
Royal Institute of Technology (KTH)

Master of Science Thesis

Frequency Analysis of Low-Frequency Field Fluctuations Detected by the Cluster Satellites

Author: Yiwen Xu

Supervisor: Prof. Lars Blomberg
Dr. Tommy Eriksson

December 18, 2007
Royal Institute of Technology Report XR-EE-SPP 2007:006
Abstract

Ultra-low frequency fluctuations in the frequency range 1 mHz to 1 Hz were already observed on ground two centuries ago. Thanks to the Cluster mission, we can now study these pulsations more closely and accurately than ever. Measurements of electric and magnetic fields in the Earth’s magnetosphere are collected by instruments onboard the four spacecraft of the Cluster mission and four sets of data from four different days are studied in this thesis. The Fourier transform, non-parametric and parametric spectral estimation methods, analytic signal, Short-Time Fourier transform (STFT) and wavelets are applied to the satellite data to analyze their frequency contents. The Fourier transform and the non-parametric spectral estimation methods offer a simple tool to gain an overall picture of the frequency contents while the analytic signal, STFT and wavelets provide a good tool for looking at the frequency as a function of time. The parametric approach in estimating power spectrum density and frequency should be avoided because the fluctuations in question can not be modeled efficiently.
Acknowledgements

I would like to thank Professor Lars Blomberg for guidance and support for this thesis. Many thanks to Dr. Tommy Eriksson who has helped me patiently with all kinds of questions and problems even after he left the lab. A big thanks to Tanja, Gael, Thu, Markus, Markus and Giorgio for the nice companionship at the lab. Last but not least, I would like to express my gratitude to my family for support and love.
## Contents

1 Introduction 5

2 Space Plasma Physics 6
   2.1 Plasma Physics ........................................... 6
   2.2 Waves in Plasma ........................................... 7
   2.3 Magnetosphere ............................................ 8
   2.4 Ultra-low Frequency Fluctuations ...................... 8

3 Cluster Mission and Instrumentation 11
   3.1 Cluster Mission ......................................... 11
   3.2 Coordinate Systems ..................................... 11
   3.3 Instrumentation ......................................... 12
      3.3.1 Satellite Data ..................................... 12
      3.3.2 Electric Field and Waves (EFW) .................. 12
      3.3.3 Fluxgate Magnetometer (FGM) .................... 13

4 Methods of Analysis 14
   4.1 Spectral Analysis ....................................... 14
      4.1.1 Fourier Transform .................................. 14
      4.1.2 Non-parametric Methods ........................... 14
      4.1.3 Parametric Methods ................................ 17
   4.2 Analytic Signal ......................................... 19
   4.3 Short-Time Fourier Transform .......................... 19
   4.4 Wavelet Analysis ........................................ 20
      4.4.1 Continuous Wavelet ................................. 21
      4.4.2 Discrete Wavelet Transform ....................... 22

5 Data Analysis and Results 27
   5.1 Simulation of test signals and satellite data ........ 27
   5.2 Analysis of satellite data .............................. 35
      5.2.1 Data Processing ..................................... 35
      5.2.2 Analysis of data from 14th of August, 2003 .... 37
      5.2.3 Analysis of data from 17th of August, 2002 ...... 38
      5.2.4 Analysis of data from 21st of September, 2002 ... 41
5.2.5 Analysis of data from 30th of August, 2001 ........................ 46

6 Conclusions and Discussions ........................................... 51

Bibliography ................................................................. 54
Chapter 1

Introduction

Ultra-low frequency fluctuations in the frequency range 1 mHz to 1 Hz were already observed on ground two centuries ago. Thanks to the Cluster mission, we can study these pulsations more closely and accurately than ever. Measurements of electric and magnetic fields in the Earth’s magnetosphere are collected by instruments onboard the four spacecrafts of the Cluster mission and four sets of data from four different days are studied in this thesis. We are particularly interested in determining and following the frequency contents of these pulsations. Besides the classic Fourier methods, this thesis also explores the use of analytic signal, Short-Time Fourier Transform and wavelets in time-frequency analysis of these data.

In this report, we will start with a brief introduction to space plasma physics to get a background on the physical properties of the ULF pulsations. Different methods used in the frequency analysis will be introduced thereafter. Then we will simulate some test signals with known frequencies and properties to evaluate these methods. Lastly, these methods are applied to satellite data, followed by a discussion and conclusion of each method.
Chapter 2

Space Plasma Physics

2.1 Plasma Physics

One amazing fact not known to many is that at least 99% of our universe is made of plasma so that it is often considered fourth state of matter. Plasma is essentially a bunch of ionized particles which has no net charge on average. The name plasma comes from the Greek word “plassein” which means “something moulded”, described by Irving Langmuir in 1929 when he saw a glowing gas fills a tube [4]. A typical plasma is a hot and highly ionized gas. Transition from the state of gas to plasma usually requires a temperature of size $10^4 – 10^5$ Kelvin and an ionization energy around $10eV$ per atom. On Earth, we are well familiar with three states of matter since most of the Earth’s near environment is not suitable for existence of plasma. Phenomena such as polar auroras and lightning are all examples of terrestrial plasma. Plasma can also be made artificially; the fluorescent lamp used in daily life is such an example.

The dynamics of plasma is rather complex; one needs to consider the interaction of charge carriers as well as their behavior in presence of the electric and magnetic fields. One simple approach is to study plasma based on the single particle motion instead of the collective behavior of plasma. The other approach is the opposite extreme case where single particle aspects are ignored completely [2]. Magnetohydrodynamics (MHD) is the field of physics that incorporates not only the mechanical laws but also electromagnetic properties in order to describe and study the behaviors of plasma dynamics. The set of equations used in MHD is a combination of Navier-Stokes equations of fluid dynamics and Maxwell’s equation of electromagnetism. This approach has also been proven quite suitable for the study of low-frequency waves in magnetic plasmas.

Apart from the two fundamental parameters density and temperature, a number of other parameters are introduced to describe the plasma. One them is the Debye length $\lambda_D$ which indicates the distance over which a plasma ion influences its surroundings. In other words, the Debye length is determined the by balance obtained between the thermal particle energy and the electrostatic potential energy. As one can see from
equation 2.1, $\lambda_D$ is dependent on the temperature $T$, the electron-number density $n$ in an electron-proton plasma. Here $k$ denotes the Boltzmann's constant and $e$ the electron charge. Plasma in the magnetosphere and solar wind has a Debye length of the scale from 1 to 100m.

$$\lambda_D = \sqrt{\epsilon_0 k T / n e^2}$$  \hspace{1cm} (2.1)

Another parameter of importance closely related to the Debye length is the plasma frequency, the natural frequency of oscillations in plasma given in equation 2.2 below. For unmagnetized plasma, the plasma frequency is the lowest frequency that an electromagnetic wave can have in order to propagate in the plasma.

$$\omega_p = \sqrt{n_s e^2 / \epsilon_0 m_e}$$ \hspace{1cm} (2.2)

A charged particle in a uniform magnetic field moves in a circle. The frequency of this circular motion is defined as the gyrofrequency, particularly used in magnetic plasma. It is given in equation 2.3 in which $m_{i,e}$ denotes the mass of ion and electron respectively and the charge $e$ positive for ions and negative for electrons. This results in that circulations have different directions for differently charged particles.

$$\omega_g = eB / m_{i,e}$$ \hspace{1cm} (2.3)

### 2.2 Waves in Plasma

When perturbations occur in a system, waves are generated. This is also the case in plasma; waves are generated when disturbances propagate in plasma. By linearization of MHD equations, two restrictions, namely the dispersion relation can be derived. They have to be satisfied in order for a wave to exist.

For cold plasma where the pressure changes can be neglected:

$$\left(\omega / k\right)^2 = v_A^2 \cos^2 \theta$$ \hspace{1cm} (2.4)

$$\left(\omega / k\right)^2 = v_A^2$$ \hspace{1cm} (2.5)

For warm plasma where the pressure changes are significant:

$$\left(\omega / k\right)^2 = v_A^2 \cos^2 \theta$$ \hspace{1cm} (2.6)

$$\left(\omega / k\right)^2 = \frac{1}{2} \sqrt{v_s^2 + v_A^2} \pm \left[(v_s^2 + v_A^2)^2 - 4v_s^2 v_A^2 \cos^2 \theta\right]^{1/2}$$ \hspace{1cm} (2.7)

$v_A$ is the Alfvén velocity defined in equation 2.8, $c_s$ the speed of sound and $\theta$ the angle between the $\mathbf{B}$-field and the wave propagation vector $\mathbf{k}$.

$$v_A = \sqrt{B^2 / \mu_0 \rho}$$ \hspace{1cm} (2.8)
One type of waves, the shear *Alfvén* wave named after the Swedish physicist Hannes *Alfvén* does not change the local plasma pressure, density or the field magnitude as it propagates. However, magnetic field lines are bent during the process. This wave behaves as a transverse wave, propagating with the phase velocity $v_{\text{Alfven}}$ along the magnetic field lines. The other type of wave identified is a compressional wave. Unlike the shear *Alfvén* wave, it causes fluctuations in the density and pressure of the plasma as well as in the field magnitude and the magnetic pressure. They can travel in any direction.

In a warm MHD plasma three different wave solutions can be derived from equations 2.4 and 2.7. One being the transverse wave mentioned before and the other two the compressional waves which carry variations in plasma and magnetic pressure as well as the variations in plasma density. These compressional waves can be further divided into as the fast and slow modes when taking the direction of propagation and phase velocity into account. Apart from the shear *Alfvén* wave, only one kind of compressional wave occurs in a cold plasma.

### 2.3 Magnetosphere

The magnetosphere is Earth’s shield against the solar wind and cosmic radiations. It is formed by the interaction of solar wind and Earth’s magnetic field. See figure 2.1 for an illustration. Despite its name, it’s no perfect sphere above the atmosphere. On the side facing the sun, the distance to its boundary is about 70,000 km and on the opposite side, the magnetic tail of the magnetosphere can stretch up to orbit of our moon [4]. The plasma in the magnetosphere consists mainly of electrons and protons originated from the solar wind and the ionosphere which is the uppermost part of the atmosphere. The solar wind is essentially hot plasma embedded with interplanetary magnetic field originating from the sun. When it meets the terrestrial plasma, a boundary known as the magnetopause is formed. The terrestrial plasma is pushed into a cavity bounded by the magnetopause called the magnetosphere. The solar wind itself is decelerated in the magnetosheath and flows around the magnetopause. The typical velocity of the solar wind is around 450 km/s while the *Alfvén* velocity and speed of sound are around 60 km/s under the same circumstances. A distinctive feature of the magnetosphere, a bow shock, is formed when the velocity of the solar wind drops from "supersonic" to "subsonic". The Earth’s bow shock is about 100-1000 km thick and located about 90,000 km from the Earth [4]. The ionosphere is located from 90 up to 1500 km above the Earth and forms the inner edge of the magnetosphere. In other words, the magnetosphere can be seen as bounded by the ionosphere and the magnetopause.

### 2.4 Ultra-low Frequency Fluctuations

A wave which has lower frequency than its natural plasma frequency and gyrofrequency is defined as ultra-low frequency (ULF) waves. This class of waves covers roughly the
frequency range from 1 mHz to 1 Hz. The ULF pulsations can then be further classified into continuous and irregular pulsations, denoted by Pc and Pi. The continuous pulsation Pc is characterized by its quasi-sinusoidal shape and well-defined spectral peaks. They can have durations from several minutes up to hours. The irregular pulsations are pulsations in the same frequency band but contain power at many different frequencies. They are usually short-lived and have a broad spectrum [2]. See table 2.1 below for durations and frequency range of the two kinds of pulsations.

<table>
<thead>
<tr>
<th></th>
<th>Pc-1</th>
<th>Pc-2</th>
<th>Pc-3</th>
<th>Pc-4</th>
<th>Pc-5</th>
<th>Pi-1</th>
<th>Pi-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>T(s)</td>
<td>0.2-5</td>
<td>5-10</td>
<td>10-45</td>
<td>45-150</td>
<td>150-600</td>
<td>1-40</td>
<td>40-150</td>
</tr>
<tr>
<td>f(Hz)</td>
<td>0.2-5</td>
<td>0.1-0.2</td>
<td>0.022-0.1</td>
<td>0.007-0.022</td>
<td>0.002-0.007</td>
<td>0.025-1</td>
<td>0.002-0.025</td>
</tr>
</tbody>
</table>

Table 2.1: Duration times and frequencies of different pulsation classes [3]

Ultra-low frequency (ULF) fluctuations of the Earth’s magnetic field were observed on ground already in the 19th century but they were not intensively studied until a century later. It was suggested by Dungey in the 1950’s that the MHD waves standing along magnetic lines and reflected at the ionosphere at the two ends were sources of these oscillations of magnetic field [3]. For standing waves to exist on the field lines, only certain resonant frequencies are allowed. The ionosphere plays an important role; it transmits and reflects ULF waves. The magnetosphere is often modeled as a box bounded by the ionosphere and the magnetopause to estimate the allowed frequencies. In many cases, the pulsations observed are associated with the Alfvén waves [2]. Three dominant ways to study the ULF fluctuations now days are ground-based magnetometers, ionospheric radars and satellites. The mechanism behind the generation of ULF pulsations has been intensively studied in recent years to deepen our knowledge about
the magnetosphere, especially in classifying Pc5 pulsations into waves with small azimuthal wave-number, $|m| \sim 1$ and large azimuthal wave-number, $|m| >> 1$. This number indicates the wave mode from which we can identify whether the driving mechanism is external or internal of the magnetosphere [1].
Chapter 3

Cluster Mission and Instrumentation

3.1 Cluster Mission

The Cluster mission was first proposed in 1982 and finally launched in 2000 by the European Space Agency (ESA). The aim of the Cluster mission is to study small-scale structures of the magnetosphere and its environment in three dimensions. To achieve this, four identical spacecraft fly in an irregular tetrahedral configuration [6]. Multi-point measurements greatly increase the amount of information that can be obtained compared to single-point measurements. The most prominent advantage is the separation between spatial and temporal variations [1]. The orbits of these satellites extend from 19 000 up to 119 000 km, moving through the solar wind and passing by the bow shock and the magnetopause. It takes approximately 57 hours for each satellite to complete its orbit. The four satellites were given the names Rumba, Salsa, Samba, Tango for publicity purposes, although most of the scientific community still lovingly refer to them as One, Two, Three, and Four [7]. Their acronyms used in this thesis are simply SC1, SC2, SC3 and SC4. The two instruments on board that we focus on here are the Electric Field and Waves (EFW) instrument and the Fluxgate magnetometer (FGM) measuring the electric and the magnetic fields in the magnetosphere.

3.2 Coordinate Systems

Below is a short description of the coordinate system mentioned in this thesis.

Geocentric Solar Ecliptic System (GSE)

X-axis: Pointing from the Earth towards the sun.
Y-axis: In the ecliptic plane pointing towards dusk. The ecliptic plane is the geometric plane containing the mean orbit of the Earth around the Sun [4].
Z-axis: Parallel to the ecliptic pole. The ecliptic poles represent the ends of an imagi-
Figure 3.1: Artistic impression of the four Cluster spacecraft.

nary line perpendicular to the ecliptic plane.

3.3 Instrumentation

3.3.1 Satellite Data

All data used in this thesis are available from the Cluster Science Data System. Since we have a rather complicated system with four satellites, many parameters have been introduced to describe their positions, velocities, orientation of the spacecraft spin axes and etc. apart from the measurements they take. Different measurements are saved in Common Data Format (CDF) files and available on daily basis.

3.3.2 Electric Field and Waves (EFW)

The Electric Field and Waves (EFW) experiment has four wire booms mounted on each of the Cluster spacecraft. The distance from the satellite center to the boom tip is 44 m and at the end of each boom, a sensor made of metal is placed [8]. In the electric field mode of this instrument, the probes are actively controlled by a bias current to ensure a high quality measurement. The electric field is measured by calculating the potential difference between the two sensors and since there are two pairs, two components of the electric field may be computed, namely the sunward and duskward components which roughly coincides with the x and y-components in GSE-coordinates respectively. The duskward electric field is computed by applying a least-squares fit of a sine wave with the spin frequency to the measured data points during one spin. However, those data have not been corrected for offsets present due to, e.g. photoelectron asymmetries [7]. For many purposes, this is acceptable; with an offset in E-field primarily in the X_GSE direction. Unfortunately, the sunward electric field is severely affected by the asymmetries and therefore it’s often left out in further studies and analysis conducted on the electric field. The CDF files that contain the EFW data are named after its corresponding satellite and date. For example, EFW data collected for SC1 on 17th of August 2002 is simply C1_PP_EFW_20020817.CDF.
3.3.3 Fluxgate Magnetometer (FGM)

The basic principle of the fluxgate magnetometer is to compare the drive-coil current needed to saturate the core of a toroidal in one direction as opposed to the opposite direction. The difference is due to the external field [9]. On each spacecraft an identical FGM instrument is placed to measure the magnetic field. Each instrument, in turn, consists of two triaxial fluxgate magnetometers and an onboard data processing unit. Three components of the magnetic field are measured and sampled at a constant rate of 201.75 Hz [7]. Sensors of the FGM are placed far from each other in order to minimize the magnetic background of the spacecraft. Similarly, measurements of the magnetic field are saved on daily basis in their corresponding CDF files.
Chapter 4

Methods of Analysis

In this study and analysis of ULF pulsations, we are particularly interested in determining the frequency contents of these pulsations and variations in frequency. We will first begin with the classic Fourier transform and spectral estimation methods to obtain information about the frequency of the pulsations. Then we will move onto the usage of analytic signal, short-time Fourier transform and wavelets to get a better picture of how pulsations vary both in frequency and time. Background and theory of these methods are presented below.

4.1 Spectral Analysis

4.1.1 Fourier Transform

The Fourier transform (FT) is a true classic for analyzing signals. The Fourier transform that we use, Discrete Fourier Transform (DFT) in equation 4.1 breaks down a signal into constituent sinusoids of different frequencies which offers a good tool to look at the same signal in the frequency domain [13]. In other words, the Fourier transform works like a frequency sorting process and ideally, only at the true frequency of the signal, would there be a peak value with zero bandwidth. However, this is seldom the case; a phenomenon known as spectral leakage introduced by sudden changes in the signal at the start and the end causes other frequencies than the true ones to also obtain a non-zero value as if the signal has leaked from its true frequency to adjacent frequencies. There are two ways to tackle the problem; the first one is to increase the length of the DFT and the other is to modify the DFT by applying windows.

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N}, \quad k = 0, 1, ..., N - 1$$

4.1.2 Non-parametric Methods

Associated with the physical background of the signal, a more meaningful alternative tool is to estimate the power spectral density (PSD), which describes how the power
of a signal is distributed over frequency. Two approaches to calculate the PSD are the non-parametric and the parametric spectral estimation methods. The non-parametric spectral methods are based on the Fourier transform of either the signal itself directly as in equation 4.2 or the Fourier transform of its autocorrelation function in equation 4.3. In many cases, the autocorrelation function of the signal is unknown and has to be estimated based on the data available, often finite and noisy using equation 4.4 and 4.5.

\[
\hat{P}_{xx}(f) = \frac{1}{N} \left| F\{x_N(n)\} \right|^2 = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n)e^{-j2\pi fn} \right|^2 \quad (4.2)
\]

\[
\hat{P}_{xx}(f) = F\{\hat{r}_{xx}(k)\} = \sum_{k=-N+1}^{N-1} \hat{r}_{xx}(k)e^{-j2\pi fn} \quad (4.3)
\]

\[
\hat{r}_{xx}(k) = \frac{1}{N} \sum_{n=0}^{N-k-1} x(n+k)x^*(n) \quad k = 0,1,\ldots,N-1 \quad (4.4)
\]

\[
\hat{r}_{xx}(k) = \hat{r}_{xx}^*(-k) \quad k = -1,-2,\ldots,-N+1 \quad (4.5)
\]

Equations 4.2 and 4.3 are foundations of the most common method used to estimate power spectrum density, the periodogram. In fact, the periodogram is equivalent to multiplying a rectangular windows of height unity to the data series as shown in equation 4.6. Although the periodogram is easy to compute, it is limited in its ability to produce an accurate estimate of the power spectrum, particularly for short data records [10]. As already mentioned before, leakage is an inevitable problem in the calculations of the DFT and since the periodogram is based on the DFT, this problem also occurs here. With the rectangular window in the periodogram, this results in the narrowest main lobe of all windows that can be applied at the expense of the largest side lobes [11]. This result is sometimes misleading if for example, a signal contains two frequencies very close to each other and only one frequency is observed while the other drowns in the large side lobes.

\[
\hat{P}_{xx}(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n)w(n)e^{-j2\pi fn} \right|^2 \quad (4.6)
\]

where \(w(n)_R\) denotes a rectangular window of height unity.

Apart from the rectangular window, other windows can also be applied to the data series which will reduce the effect of leakage to different degrees. Examples are a Bartlett window with a triangular shape or a Hamming window displayed in figure 4.1. Windows are not magic; what they provide is a trade-off between spectral resolution (main lobe width) and the spectral masking (side lobe amplitude) [10]. The edge of the window causes information loss at the endpoints; shape edges will lead to large side

15
lobes.

In this thesis, the two methods mainly used are `pwelch` and `pmtm` functions in Matlab. For the `pwelch` function in Matlab, data are divided into eight segments with 50% overlap. Each segment is windowed with a Hamming window and eight modified periodograms are computed and averaged, mathematically shown in equation 4.7. Theoretically, the Hamming window is especially useful in those applications that can tolerate very little leakage. The Hamming window generates very small side lobes, but results in a main lobe that is twice as wide as that of the rectangular window. Therefore, the Hamming window causes a great deal of local spectral leakage but very little distant leakage [11]. The fact that the data have been divided into segments with overlaps and averaged is an effort to reduce variance.

\[
\hat{P}_{xx}(f) = \frac{1}{K} \sum_{k=0}^{K-1} \frac{1}{L} \sum_{n=0}^{L-1} x_k(n)w(n)_He^{-j2\pi fn} \quad (4.7)
\]

here \( K \) is the number of segments, \( L \) the length of each segment and the normalization constant \( U = \frac{1}{L} \sum_{n=0}^{L-1} \left| w(n)_H \right|^2 \).

The `pmtm` method is based on Thomson’s multi-taper method (MTM) which uses linear or nonlinear combinations of modified periodograms to estimate the PSD. These periodograms are computed using a sequence of orthogonal tapers (windows in the frequency domain) specified from the discrete prolate spheroidal sequences (DPSS) [13]. An illustration of DPSS windows is displayed in figure 4.2. There is a whole family of tapered windows and they all have the feature of gradually tapering the data near the end points of the window, thus leading to improvements in narrowing the main lobe with some reduction in the size of the side lobes. The more tapers one uses, the wider the main lobe [11].

Figure 4.1: Different windows: Bartlett, Blackman, Hamming and Gaussian window.
4.1.3 Parametric Methods

In some applications, however, estimating the power spectrum may be facilitated by having prior knowledge about how the process is generated. This type of information may then allow one to parametrically estimate the power spectrum known as the parametric method [10]. In a nutshell, as illustrated in figure 4.3, the parametric method tries to model a filter with frequency response $H(f)$ so that when a white noise $e(n)$ with unit variance passes through the filter, the output generated $\hat{y}(n)$ is as close to the original signal $y(n)$ as possible. Then the PSD is estimated with the help of equation 4.8, where $\hat{\sigma}_e$ is the variance of the white noise $e(n)$. Alternatively, we can rewrite this equation by specifying the frequency response as in equation 4.9 from which we can define different types of models. Here $b_q(k)$ and $a_p(k)$ denote the model parameters and $p$ and $q$ are the order of the model. More specifically, this is in fact the autoregressive moving average spectrum estimation (ARMA). If only the $b_q(k)$ coefficients are calculated, it is the moving average (MA) model and with only $a_p(k)$, it is known as the autoregressive (AR) model.

$$\hat{P}_{xx}(f) = \hat{\sigma}_e^2 | \hat{H}(f) |^2$$

$$\hat{P}_{xx}(f) = | H(f) |^2 = \frac{\left| \sum_{k=0}^{q} \hat{b}_k e^{-j2\pi fk} \right|^2}{1 + \sum_{k=1}^{p} \hat{a}_k e^{-j2\pi fk}}$$

One of the parametric methods used in this thesis is the Burg’s method, implemented with the `pburg` function. An autoregressive (AR) prediction model for the signal is calculated based on the criteria that the sum of squares of the forward and
backward prediction error is minimized [13]. Solving the autocorrelation normal equation, we obtain the AR coefficients through a method called the Yule-Walker method. It fits an autoregressive (AR) linear prediction filter model to the signal by minimizing the forward prediction error in the least squares sense. The corresponding function for this method is *pyule* in Matlab. In a nutshell, an AR model of order $p$ models the time series as autoregressive, which means that the sample in consideration is a linear function of the preceding $p$ samples. Therefore it is appropriate for a linear dynamical system which can be completely described by a set of eigenmodes with well defined frequencies where energy is concentrated in sharp peaks. For the type of problems facing the physicist studying processes in a turbulent space plasma, this is a rare case [12]. Nevertheless, it would be interesting to see how they perform on ULF pulsations.

Frequency estimation methods also belong to the parametric family based on the assumption that process under study is composed of complex frequency components in the presence of noise as in equation 4.10. Instead of estimating the power spectrum density, the frequency of the signal is estimated directly. This type of methods is implemented from the eigendecomposition of the autocorrelation matrix into two subspaces namely the signal and noise subspace. If the eigenvalues are sorted in decreasing order, the $p$ eigenvectors corresponding to the $p$ largest eigenvalues span the signal subspace. The remaining eigenvectors span the noise subspace. A frequency estimation function can be defined as in equation 4.11 and 4.12. In theory, it would be infinitely large if $e$ is orthogonal to $v_i$ at the frequencies when $\omega = \omega_i$ [10]. From the sharp peaks of the frequency estimation function, frequencies are evaluated. In this study, since the pulsations are embedded in noise, two different algorithms, namely the multiple signal classification method (MUSIC) and the eigenvector method are tested on satellite data to estimate the frequency and power content. The corresponding functions are *rootmusic* and *rooteig* in matlab. Both methods are most useful for frequency estimation of signals made up of a sum of sinusoids embedded in additive white Gaussian noise [13].

$$x(n) = \sum_{i=1}^{p} A_i e^{j\omega_i n} + e(n)$$  \hspace{1cm} (4.10)
MUSIC method:

\[ \hat{P}_{MU}(e^{j\omega}) = \frac{1}{\sum_{i=p+1}^{M} |e^{iH\mathbf{v}_i}|^2} \]  \hspace{1cm} (4.11)

Eigenvector method:

\[ \hat{P}_{EV}(e^{j\omega}) = \frac{1}{\sum_{i=p+1}^{M} \frac{1}{\lambda_i} |e^{iH\mathbf{v}_i}|^2} \]  \hspace{1cm} (4.12)

where \( \mathbf{v}_i \) are the noise eigenvectors, \( \mathbf{e} = [1, e^{j\omega_1}, e^{j2\omega_1}, \ldots, e^{j(M-1)\omega_1}] \) and \( \lambda_i \) is the eigenvalue associated with the eigenvector \( \mathbf{v}_i \). The symbol \( H \) denotes the Hermitian transpose.

4.2 Analytic Signal

A signal can be represented in its analytic form which only deals with its positive frequency components. This is known as the analytic signal defined in equation 4.14 which is essentially the signal itself as the real part and the Hilbert transform of the signal as its imaginary part. A Hilbert transform is an operation that represents a time-domain signal that is 90 degrees "out of phase" with the input signal. Another way of representing the analytic signal since it's a complex number is to use its polar form as in equation 4.15. From there, we can define some useful terms. The amplitude of equation 4.15 is called the amplitude envelope and just as the name indicates, it wraps around the signal. The argument of equation 4.15 is the instantaneous phase of the signal. The derivative of the instantaneous phase is the instantaneous frequency of the signal which will be used extensively in this study. The usage of analytic signal is constrained to quasi-monochromatic signals to obtain reasonable results. The more closely a signal approaches a narrowband condition, the better the Hilbert transform approximates the quadrature signal, and the more likely the Hilbert transform-based analytic signal is to provide an accurate model of a real system with a particular instantaneous frequency; also the better in general will be the estimate of instantaneous frequency [14].

Hilbert Transform:

\[ F_{Hi}(t) = -\frac{1}{\pi t} * x(t) \]  \hspace{1cm} (4.13)

Analytic Representation:

\[ X(t) = x(t) + iF_{Hi}(t) \]  \hspace{1cm} (4.14)

\[ X(t) = A(t)e^{j\phi(t)} \]  \hspace{1cm} (4.15)

Instantaneous Frequency:

\[ f(t) = \frac{1}{2\pi} \frac{\partial}{\partial t} \phi(t) \]  \hspace{1cm} (4.16)

4.3 Short-Time Fourier Transform

When one chooses to Fourier transform a signal, the frequency contents of the signal are revealed but the information about time is lost. This is the biggest drawback of
using the Fourier methods. The ULF pulsations are subjected to a lot of surrounding disturbances and the measurements made by the spacecraft are a function of both space and time. In addition, interest is often directed toward boundary layers or other inhomogeneous regions, where the record of data will be non-stationary, due to the motion of the spacecraft through the inhomogeneous medium, even if it results from processes stationary at any single point in the plasma [12]. Therefore, a better method to study those signals is needed to keep track of changes both over frequency and time.

An improved version of the Fourier methods was first adapted by Dennis Gabor in the 1940’s who used a window to study a small section of the whole signal which can be assumed to be stationary in the width of the window. This method was named Short-Time Fourier Transform (STFT) illustrated in equation 4.17. The signal $x(n)$ is mapped onto the frequency domain by the Fourier transform and information in the time domain is obtained by shifting the window $w(n)$ to different times in steps of $u$. This is known as time-frequency analysis. The matlab function Spectrogram is an application of the STFT in which a Hamming window and overlap of the data segments are used. The spectrogram can be interpreted as an energy density denoted as $P_S$ in equation 4.18. The spectrograms used here are calculated using a window size corresponding to 350 seconds of data with an overlap of 340 seconds and plotted using logarithmic scales.

A trade-off situation between frequency and time is expressed in equation 4.19. By having small windows, the time resolution is increased at the expense of decreasing frequency resolution. On the contrary, if the windows are made large, the frequency resolution is increased while the time resolution is deteriorated. This is known as Heisenberg’s uncertainty principle. No matter the trade-off, the size of the window is constant over time for STFT which contributes to the greatest shortcoming of this method, making it not very practical and suitable for analyzing signals that change rapidly and require a high resolution at certain frequencies.

$$X(u, f) = \sum_{n=-\infty}^{\infty} x(n)w(n-u)e^{-j2\pi fn}$$  \hspace{1cm} (4.17)

$$P_S(u, f) = |X(u, f)|^2 = \sum_{n=-\infty}^{\infty} x(n)w(n-u)e^{-j2\pi fn} |^2$$ \hspace{1cm} (4.18)

$$\triangle f \triangle t = 1$$ \hspace{1cm} (4.19)

4.4 Wavelet Analysis

As remarked earlier, one serious limitation of the STFT is that it uses a window of fixed size on all data. This is how wavelets come into the picture, offering us a flexible size of the windowing depending on the context. Fourier analysis consists of breaking up a signal into sine waves of various frequencies. Similarly, the wavelet analysis is the breaking up of a signal into shifted and scaled versions of the mother wavelet. In other words,
wavelets are local in both frequency /scale (via dilations) and in time (via translations). One major advantage afforded by wavelets is the ability to perform local analysis, that is, to analyze a localized area of a larger signal. For non-periodic and/or non-stationary signals, it's preferred to use wavelet rather than traditional Fourier methods [13].

4.4.1 Continuous Wavelet

The word wavelet is due to Morlet and Grossmann in the early 1980s. They used the French word ondelette, meaning "small wave" [4]. Like its name, wavelets are small waves with limited duration and their value is zero on average. There is a whole family of wavelets but the Morlet wavelet is most commonly used in space plasma applications with its clear physical interpretation as a modulated sinusoidal oscillation and good properties of localization in frequency as well as in time [12]. Equation 4.20 is known as the mother Morlet wavelet and by introducing parameters $s$ and $u$ as the scale and the translation respectively, we get an expression for the daughter wavelet in equation 4.21. The term $\frac{1}{\sqrt{s}}$ is introduced to ensure that the wavelet has the same energy for all scales. In figure 4.4, we can see a picture of different daughter wavelets obtained by changing the scale $s$ and the translation parameter $u$ of the mother wavelet. The term frequency is reserved for Fourier transform and scales are used instead in wavelets. In short, scale is the inverse of frequency with a normalization factor for different wavelet. In the case of Morlet wavelet, frequency $f$ is related to scale $s$ through the relationship $f = \frac{5}{2\pi s}$.

The continuous wavelet transform (CWT) is thus defined similarly as the Fourier transform in equation 4.22, projecting the signal onto scaled and shifted version of the mother wavelet to calculate wavelet coefficients which measures the variation of frequency in a neighborhood of $u$ whose size is proportional to $s$. Sharp signal transitions create large amplitude wavelet coefficients which make wavelets a good tool in analyzing trends, breakdown points and discontinuities in signals. Similarly as in the STFT, we can define a corresponding energy density function by the name scalogram expressed in equation 4.23. The scalograms used here are calculated using the Morlet wavelet and plotted in logarithmic scales.

$$\psi(t) = e^{-\frac{t^2}{2}} \cos(5t)$$  \hspace{1cm} (4.20)

$$\psi_{us}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right)$$  \hspace{1cm} (4.21)

$$C(u, s) = \int x(t)\psi^*_{us}(t) \, dt$$  \hspace{1cm} (4.22)

where * denotes the complex conjugate of $\psi$.

$$P_W(u, s) = |C(u, s)|^2 = |\int x(t)\psi^*_{us}(t) \, dt|^2$$  \hspace{1cm} (4.23)
Wavelet transforms enable us to obtain orthonormal basis expansions of a signal using time-frequency atoms that have good properties of localization in the time and frequency domains and this basic idea as well as Heisenberg's uncertainty principle are illustrated in figure 4.5. In any such decomposition the time-frequency plane is layered with cells, called Heisenberg cells, whose minimum area is determined by the uncertainty principle [19]. In the first subfigure, we can see that using standard basis such as Diract delta function, we have a high resolution in time. For the Fourier basis, all the information is located in the frequency domain as mentioned previously. With windowed Fourier basis, namely the STFT, the resolution is improved in the time-frequency domain but fixed for all time and frequencies. Lastly, the wavelet transform is based on an octave band decomposition of the time-frequency plane. From the different size of boxes, we observe that good time resolution and poor frequency resolution occur at high frequencies, while at low frequencies, we get a better resolution in frequency but less satisfactory in time. There is no significant difference due to the choice of wavelets and they pick up similar information subject to the constrain of resolution of Heisenberg cells [19]. Since we are dealing with ultra-low frequency pulsations, theoretically, we should get a very good resolution in terms of the frequency using wavelet analysis.

4.4.2 Discrete Wavelet Transform

The corresponding transform for discrete signals is the discrete wavelet transform which is implemented on discrete values of scale and translation illustrated in equation 4.24. The discrete wavelet transform is given below in equation 4.25. Here $s$ is chosen as
Figure 4.5: Schematic of time-frequency plane decomposition using different bases: (a) Standard basis, (b) Fourier basis, (c) windowed Fourier basis, and (d) wavelet basis [19].

\[ s = s_0^j \] where \( j \) is an integer and \( s_0 \) is a fixed dilation step greater than 1. Furthermore, \( t \) is chosen so that \( t = ku_0 s_0^j \), where \( u_0 > 0 \) depends upon \( \psi(t) \) and \( k \) is an integer [19]. The choices of \( s_0 \) and \( u_0 \) depend on the type of wavelet used. If \( s_0 \) is set to 2 and \( u_0 \) to 1, then we obtain the dyadic scales and it can be shown that our analysis will be much more efficient and just as accurate as in the continuous case [13]. One thing worth mentioning here is that the discrete wavelet transform is in fact not a real discrete transform since only the scales and translations are discrete. It has been proven that the necessary and sufficient conditions for stable reconstruction of original signal is that the energy of the wavelet coefficients must lie between two positive bounds [22]. In practice, the original signal can be reconstructed by summing the orthogonal wavelet basis functions, weighted by the wavelet transform coefficients.

\[
\psi_{j,k}(t) = s_0^{-j/2} \psi(s_0^{-j} t - ku_0) 
\]

\[
C(j,k) = s_0^{-j/2} \int x(t) \psi(s_0^{-j} t - ku_0) \, dt 
\]

We are particularly interested in the orthogonal wavelet transform due to its applications in the multi-resolution analysis which allows us to decompose a signal into several orthogonal time-dependent components with different frequency scales. We can find an orthogonal wavelet basis \( \psi_{j,k} \) for \( s_0 = 2 \) and \( u_0 = 1 \) such that the set of functions.
\{\psi_{j,k}\} for all \(j\) and \(k\) form an orthonormal basis in equation 4.26. The most remarkable property of this basis is that the functions are orthogonal to their translates and dilates. For practical applications, often a smoother basis is desired [19]. One example of such wavelet was developed by Daubechies called D8. The wavelet and its Fourier transform are displayed at the top of figure 4.6. Two terms often used in the context of multi-resolution analysis are the approximations and details of a signal. Approximations \(a\) are the low frequency components and details \(d\) represent the high frequency components. Another way of expressing the discrete wavelet transform is the wavelet analysis and synthesis form of the order \(J\) in equation 4.27 and 4.28 and the signal is simply the sum of all the approximations and details thanks to fact that the bases are all orthogonal. This is also known as wavelet decomposition and will be explored further later.

\[
\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - k2^j}{2^j}\right) \quad (4.26)
\]

Analysis:

\[
C(j, k) = \int x(t) \psi_{j,k}(t) \quad (4.27)
\]

Synthesis:

\[
x(t) = \sum_j \sum_k C(j, k) \phi_{j,k}(t) = a_j + \sum_{j \leq J} d_j = a_j + \sum_{j \leq J} \sum_k C(j, k) \phi_{j,k}(t) \quad (4.28)
\]

where \(s_0 = 2^j\) and \(u_0 = k2^j\), and function \(d_j(t)\) represents wavelet details of order \(j\) and \(a_J\) the approximation of signal \(x(t)\) [20].

Figure 4.6: Daubechies D8 wavelet and its Fourier transform magnitude in the upper panel and in the bottom panel the corresponding scaling function and its Fourier transform magnitude [19].
The continuous wavelet transform maps a one dimensional signal to a two dimensional time-scale joint representation that is highly redundant. The introduction of the discrete wavelet transform removes the redundancy issue and in order to limit the number of scales and translations in the actual implementation which otherwise would be infinite, scaling functions defined in equation 4.29 are introduced. Refer to the bottom panel of figure 4.6 for an illustration of Daubechies D8 wavelet’s corresponding scaling function and its Fourier transform. Furthermore, the two-scale equations are introduced to relate the scaling function and wavelet function at two different time scales from which two filters can be derived. In summary, the wavelet can be seen as a band-pass filter and its scaling function as a low-pass filter, then a series of dilated wavelets together with a scaling function can be seen as a filter bank [22]. This is how the “real” discrete wavelet transform (DWT) is implemented in practice with signals passing through a filter bank with filter coefficients $h(n)$ and $g(n)$ illustrated in figure 4.7.

$$w_{j,k}(t) = \frac{1}{\sqrt{2^j}}w\left(\frac{t - k2^j}{2^j}\right) \quad (4.29)$$

The Two-Scale equations:

$$\frac{1}{\sqrt{2}}w\left(\frac{t}{2}\right) = \sum_n h(n)w(t - n) \quad (4.30)$$

$$\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) = \sum_n g(n)w(t - n) \quad (4.31)$$

The signal $a_j$ is filtered with a low pass filter $h$ to analyze its low frequency components and through a high pass filter $g$ for its high frequency components. The resolution of the signal, which is a measure of the amount of detail information in the signal, is changed by the filtering operations, and the scale is changed by upsampling and downsampling operations [21]. Upsampling means that we insert samples in-between the original samples while the downsampling operation picks up the certain samples of the original samples.

The consequences of the signal passing through the low pass filter and the high pass filter respectively, is that only the frequency components lower or higher than the cutoff frequency remain. This is then downsampled by a factor of 2. The output from the high pass filter is denoted as $d_{j+1}$ representing the details coefficients. From the low pass filter after downsampling we obtain the approximation coefficients $a_{j+1}$. Together, they make up the discrete wavelet coefficients. The reason behind downsampling is the Nyquist sampling theory. The original signal has its highest frequency of $\omega$ which requires a sampling rate of $2\omega$. At the output of the each filter, the remaining two pieces of the signal have $\omega/2$ as its highest frequency component and therefore we can downsample the signal and still retain sufficient information of the signal. At this stage the signal only has a length half of its original length after having gone through the low and respective high pass filters and downsampling. The time resolution is halved after the filtering operation for the reason that only half number of samples is left to represent.
the signal. The scale of the signal on the other hand is now doubled seeing that the uncertainty in frequency is reduced when only low or high frequency components are represented after the operation. This procedure is illustrated at top of figure 4.7 and mathematically summarized in equations 4.32 and 4.33.

\[
a_{j+1}(k) = \sum_{n=-\infty}^{\infty} a_j(n) \tilde{h}(n - 2k)
\]

\[
d_{j+1}(k) = \sum_{n=-\infty}^{\infty} a_j(n) \tilde{g}(n - 2k)
\]

Figure 4.7: An illustration of discrete wavelet decomposition and reconstruction [17].

The approximation coefficients \(a_{j+1}\) is analyzed further, passing through the same low and high pass filter. This procedure is repeated until the desired level is reached. In general, the DWT offers a good time resolution at high frequencies, and good frequency resolution at low frequencies. The low and high pass filters are chosen as Quadrature Mirror Filters related by equations 4.34, 4.35 and 4.36. Reconstruction from the wavelet coefficients to the original signal is basically the reverse of the decomposition operation with first upsampling and then through another low/ high pass filter, namely \(g\) and \(h\) as illustrated at the bottom of figure 4.7. It is possible to get perfect reconstruction with a delay \(L\) if all the equations below are satisfied [15]. The wavelet used in decomposition and reconstruction in this thesis is the D5 wavelet.

\[
H(z) = \tilde{G}(-z)
\]

\[
G(z) = -\tilde{H}(-z)
\]

\[
H(z)\tilde{H}(z) + G(z)\tilde{G}(z) = 2z^{-L}
\]
Chapter 5

Data Analysis and Results

5.1 Simulation of test signals and satellite data

Before applying the frequency analysis methods mentioned in the previous chapter to the satellite data, we want to test them first with some simulated data with known frequencies and properties in order to get a better picture of the characteristics and performance of these methods.

We will start with a simple mono-component non-stationary signal with a linear frequency modulation that increases from 1 to 20 mHz. The signal is sampled every 4 seconds and simulated over a two-hour interval plotted in figure 5.1. It is analyzed first with the help of the analytic signal to obtain its instantaneous frequency as function of time. It's very clear from the bottom of figure 5.1 that the frequency of the test signal varies linearly from 1 to 20 mHz. Towards the end of the signal, large oscillations and transients occur, giving us an unclear picture.

Looking at the Fourier transform of the signal in figure 5.2, we suspect that there are quite a number of frequencies existing between 2 and 18 mHz but we can not tell how the frequency is varying or if it's changing at all. In fact, the Fourier transform of another mono-component non-stationary signal with a linear frequency modulation which decreases from 20 to 1 mHz would give us the same Fourier transform. The spectral estimation methods shown at the bottom of figure 5.2 provide no more information than the Fourier transform i.e. frequencies between 2 and 18 mHz are manifested. In fact, the power spectrum density calculated with the Burg's method is misleading; five distinctive peaks are shown which makes one draw the conclusion that this signal is a multi-component signal with frequencies at 2, 5, 10, 15 and 18 mHz. This confirms the fact that AR estimators try to represent all spectral features by narrow peaks [12]. Studying the spectrogram, a clear linear growth of frequency as a function of time can be spotted in figure 5.3. In the scalogram, frequencies higher than 15 mHz is less clear than the lower frequencies. This is due to the fact that for wavelets, higher frequencies have a better resolution in time than in frequency which explains the fuzzy part in the figure.
The instantaneous frequency of the test signal

Figure 5.1: Mono-component non-stationary test signal with a linear frequency modulation and its instantaneous frequency.

Fourier transform of the signal

Figure 5.2: Fourier transform and PSD estimation of the mono-component non-stationary test signal with linear frequency modulation.

The second signal to be investigated is a multi-component stationary signal, a sum of two sinusoids with frequencies at 4 and 10 mHz. As before, the signal is sampled
every 4 seconds and over a two-hour interval plotted in figure 5.4. A plot of the instantaneous frequency obtained from the analytic signal is shown in the same figure and it is obvious that the frequency is changing but no more information is revealed. It was mentioned before that the effectiveness of using analytic signal to obtain the instantaneous frequency relies on the fact that the signal analyzed is a monochromatic signal or contains only a small band of frequencies. Filtering is suggested prior to transformation of the signal into its analytic form. This has been done for both 4 and 10 mHz to isolate them from each other and the result is plotted in figure 5.5. Now it is clear and inarguable that these two frequencies exist at all times in the signal.

The Fourier transform in figure 5.6 provides us quite accurate information about the two frequencies with some leakages. All the spectral methods seem to work quite well; they exhibit some characteristics remarked earlier. The Welch method using a Hamming window gives us very small side lobes, but results in a main lobe that is many times as wide as the periodogram with the uniform window. The Thomson multi-taper method has a smaller main lobe than the Welch method. The parametric methods seem to favor the high frequency components. Moving onto the frequency-time methods displayed in figure 5.7, we see that the spectrogram provides us good information about the frequency components of the signal as well as its localization in time which in this case is constant. The scalogram on the other hand only provides a clear picture of the 4 mHz component at the expense of less accurate time localization. Around 10 mHz it is better localized in the time domain but it is hard to tell what the exact frequency is. The frequency estimation methods are very accurate in determining
the frequencies involved; both methods have estimated exactly the two frequencies at 4 and 10 mHz. Furthermore, a wavelet decomposition of level 4 with Daubechies D5 wavelet is applied to the signal with results presented in figure 5.8. By visual inspection, it is obvious that the a4 coincides with the 4 mHz component and d4 with the 10 mHz component. The shape of the 4 and 10 mHz components has been slightly modified but it does not affect the frequency contents.
Figure 5.6: Fourier transform and the PSD estimation of the multi-component stationary test signal.

Figure 5.7: Spectrogram and scalogram of the multi-component stationary test signal.

In the real world, we seldom encounter signals without disturbances. A white noise of amplitude 0.5 is added to this signal as an experiment to determine how well the different methods perform in the presence of noise. The corresponding Fourier transform and PSD as well as the spectrogram and scalogram are shown in figure 5.9 and 5.10. More leakage is seen in the Fourier transform; however, it does not affect the two peaks
at 4 and 10 mHz significantly. The parametric methods start to fail to distinguish the two frequency components with the same order of model and the model order $p$ needs to be increased to 30 to get the same performance as before for the signal without noise. One reason behind this could be that the AR estimators are trying to represent all the frequencies introduced by noise which spreads out its spectrum to all frequencies. In the spectrogram, the energy level at other frequencies has increased due to the presence of noise and the two straight lines at 4 and 10 mHz have become blurrier. For the scalogram, we see that it’s less sensitive to noise than the spectrogram and performs roughly the same as before. To our disappointment, the frequency estimation methods have estimated frequencies at 7.6 and 37.9 mHz for MUSIC method and for eigenvector method, the approximations are 7.7 and 42.8 mHz. The four level wavelet decomposition of this noisy signal is plotted in figure 5.11. The details $d_1$, $d_2$ and $d_3$
consist mostly of noise. The 4 mHz component appears in the approximations level 4 $a_4$ and the 10mHz component at details $d_4$ just as in the case without noise. This might be a good method to detect the frequency components of a signal despite of the noise.

![Fourier transform of the signal](image1)

![Power Spectrum Density of the signal estimated with different methods](image2)

**Figure 5.9:** Fourier transform and the PSD estimation of the multi-component stationary test signal embedded in white noise

![Spectrogram for the test signal](image3)

![Scalogram for the test signal](image4)

**Figure 5.10:** Spectrogram and scalogram of the multi-component stationary test signal embedded in noise.
Figure 5.11: Wavelet decomposition of the multi-component stationary test signal embedded in noise.

Both the electric and magnetic field data are a function of space and time as well as have been subjected to a number of disturbances from various instruments and environment. The third test signal is modeled with the following parameters and assumptions, trying to mimic the behavior of real signals picked up by satellites in addition to some extreme characteristics. We simulate a signal of two-hour duration with 4 second sampling time which gives us 1800 samples in total. For the first 800 samples, we modulate a sinusoidal signal with its amplitude varying according to a Gaussian distribution. Its frequency is set to 4 mHz. Then for the next 400 samples, we set its amplitude ten times smaller than in the previous case. For the remaining 600 samples, we modulate its shape as a Gaussian curve with amplitude slightly higher than in the first 800 samples. However, the frequency is varying linearly with time from 1 to 8 mHz in the last 200 samples. Lastly, we add a white noise of amplitude 0.5 to the whole signal and the final
The instantaneous frequency of the test signal

Figure 5.12: Simulation of signal from satellites.

The instantaneous frequency plot at the bottom of figure 5.12 reveals a signal with 4 mHz frequency. Abrupt changes occur at time instants when the signal is changing significantly for example at 3000 s and 4700 s when the amplitude changes. This is an indication of the sensitivity and effectiveness of this method. The frequency variations in the last 200 samples are not detected at all. In figure 5.13, we realize that the Fourier transform and PSD estimates reveal very well the 4 mHz component but still, the linear frequency change at the end of the signal is not indicated. Some indications of the change in frequency toward the end can be seen in the spectrogram as well as in the scalogram presented in figure 5.14. However, this information is not enough for us to identify the nature of the signal in the last 200 samples due to its short period and because its frequency is changing too rapidly.

5.2 Analysis of satellite data

5.2.1 Data Processing

Data collected by the four spacecraft of the Cluster mission are far from perfect. Collection of measurements of the electric field should be sampled every 4 seconds. However, in reality, some data are missing between the samples and they are linearly interpolated before further processing. In order to compare the electric and magnetic fields, the magnetic fields are resampled to the same time scale as the electric field. The measured magnetic field has three components which consist mostly of the background field.
Figure 5.13: Fourier transform and the PSD plot of the simulated satellite signal.

Figure 5.14: Spectrogram and scalogram of the simulated satellite signal.

provided by the Earth’s magnetic field and it’s difficult to get any information about the ULF pulsations. Therefore, these data are filtered with a band pass filter to obtain the magnetic field of the ULF pulsations of the order nT.
5.2.2 Analysis of data from 14th of August, 2003

The ULF pulsations occur at about 10:30 universal time (UT) when the spacecraft are relatively close to each other. Thus the measurements of the duskward electric fields are rather similar from all spacecraft shown in figure 5.15. This event ends at around 12:30 and measurements from SC2 are further studied and analyzed. Just with the naked eye, we suspect that this is a monochromatic signal and observe around 11 cycles in half an hour interval which gives a frequency of 6 mHz.

![The duskward electric fields 2003-08-14](image)

Figure 5.15: The duskward electric field measured with Cluster satellites 2003-08-14.

Analyzing the signal with the FT and PSD, displayed in figure 5.16, gives us confirmation about the frequency of the pulsation, a clear peak shown by all methods at 5.6 mHz. Thereafter, a low-pass filter with cutoff frequency at 8 mHz is applied to the measurements and plotted with the original measurements in figure 5.17. In the spectrogram, a dominant pulsation of approximately 6 mHz occurs between 10:30 and 11:10 and also between 11:30 and 12:00. This is in agreement with measurements of electric field that ULF pulsations are most evident in those intervals. Short pulsations with other frequencies are also present, particularly at around 10:20 and 12:15. More or less the same information is obtained with the scalogram displayed in figure 5.18. The plot of the pulsation’s instantaneous frequency in figure 5.19 portrays that the frequency is oscillating around 5 mHz, consistent with the frequency estimated by other methods. The changes in frequency are most abrupt at around 10:20 and 12:15 which is where we observe the start of dominant pulsations. The estimated frequency turns out to be 9.6 mHz with the MUSIC method which occurs occasionally at 10:20 and 12:15. The other frequency estimator, namely the eigenvector method detects the most dominant frequency at 4.9 mHz.
Figure 5.16: The FT and PSD plot of duskward electric field SC2 2003-08-14.

Figure 5.17: Duskward electric field SC2 2003-08-14.

5.2.3 Analysis of data from 17th of August, 2002

On 17th of August 2002, between 16:00 and 19:00, the Cluster satellites were in the dayside magnetosphere close to perigee ($4R_E$). Between 16:00 and 18:30 ULF pulsations are clearly visible in the duskward electric field component on all four spacecraft. The oscillations are monochromatic on all four satellites and stable over both space and time.
Figure 5.18: Spectrogram and scalogram of duskward electric field SC2 2003-08-14.

Figure 5.19: Instantaneous frequency of duskward electric field SC2 2003-08-14.

[1]. As before, the duskward electric field measurements from one spacecraft, SC2 is studied in depth and plotted in figure 5.20. We observe 8 periods of the pulsation in half an hour which converts to 4.4 mHz in frequency. Seeking for confirmation in the Fourier transform and the PSD estimates in figure 5.21, a peak is apparent at approximately 4.9 mHz. However, the FT and all the PSD estimates seem to spread out in frequency especially in the interval from 3 and 6 mHz. Instead we see a sharp and huge peak at
around 0.1 mHz which could be a background field or a disturbance. The parametric methods seem to be in favor of the disturbance and completely fail to detect the 4 mHz frequency component. This is also seen earlier in the 2nd test signal that they tend to favor one frequency over another. In figure 5.22, a strong pulsation at approximately 5 mHz is visible in both the spectrogram and the scalogram from 16:30 to 17:45 and its energy decreases from then on. The extremely low frequency disturbance observed in the FT graph is also evident, particularly at 16:30 and 18:30. In figure 5.23, the graph of the instantaneous frequency, it is obvious that the frequency of the pulsation oscillates around 4.5 mHz with a deep drop in frequency at about 16:45 also observed in figure 5.22. The frequency obtained by the frequency estimators is around 12 mHz which corresponds to the observed energetic pulsations in the beginning and end of the data.

![Duskward electric field 2002−08−17 SC2](image1)

![Duskward electric field 2002−08−17 SC2](image2)

Figure 5.20: Duskward electric field SC2 2002-08-17.

It would be interesting to look at the magnetic field as well and theoretically, they should exhibit the same frequency as their electric counterparts. The three components of the magnetic field collected by SC2 are plotted in figure 5.24. The most substantial oscillations can be found in all components at about 17:30. In the FT and PSD estimates, the three most significant frequencies observed are approximately 2, 4 and 6 mHz in all components illustrated in figure 5.25. The spectrogram and the scalogram for each component of the magnetic field is plotted in figure 5.26. In the first and the third component, it’s apparent that the frequency with the highest energy is around 4.5 mHz, consistent the frequency of the electric field. The 2 and 6 mHz frequency components detected by the other methods are also visible, mostly concentrated at the end of the event and even afterwards and therefore we can disregard them. Even the frequency estimation methods give estimates of 4.6 and 5.1 mHz for the first component.
Figure 5.21: The FT and PSD plot of duskward electric field SC2 2002-08-17.

Figure 5.22: Spectrogram and scalogram of duskward electric field SC2 2002-08-17.

of the magnetic field.

5.2.4 Analysis of data from 21st of September, 2002
The ULF oscillations in the duskward electric field component start at about 11:55 on all four satellites due to their relative proximity [1]. Measurements from the 2nd
The instantaneous frequency of the duskward electric field 2002–08–17 SC2

The magnetic field component 1 2002–08–17 SC2

The magnetic field component 2 2002–08–17 SC2

The magnetic field component 3 2002–08–17 SC2

Figure 5.23: Instantaneous frequency of duskward electric field SC2 2002-08-17.

Figure 5.24: Magnetic fields SC2 2002-08-17.

satellite SC2 is chosen to be studied further and illustrated here at the top of figure 5.27. Just with the naked eye, one can suspect that there are two fluctuations, one with lower frequency at around 12:00 and another one at around 13:00 with higher frequency. This is further confirmed in the FT and PSD plot in figure 5.28. One peak occurs at around 5 mHz and a peak with amplitude half of the previous one occurs at approximately 27 mHz. There are also some smaller peaks sticking out in the interval spanned
by these two frequencies and it is difficult to tell if they are actually ULF pulsations, especially the one at around 14 mHz which is only detected by the Fourier transform and Thomson Multi-taper method. In figure 5.29, a strong pulsation at about 6 mHz is visible from 11:45 to 12:15 in the spectrogram and the same information is obtained in the scalogram. At around 13:00, a pulsation around 28 mHz appears, lasting about 15 minutes in the spectrogram. However in the scalogram, the frequency range at which this pulsation occurs spans from 20 mHz to 40 mHz, making it difficult to pinpoint the exact frequency or at least to the same extent as in the spectrogram. This owes to the fact that was remarked earlier that in wavelet analysis, for high frequencies, the time resolution is good while the frequency resolution is poor. A weaker pulsation is conspicuous at 13:20 at around 15 mHz, corresponds to one of the “suspicious” peaks displayed in the FT and PSD plot. This 15 mHz pulsation has actually occurred very shortly already at 11:35 and it’s only observed with the scalogram. In the scalogram, the pulsation with frequency of around 28 mHz is seen from 11:40 onwards but it is less obvious in the spectrogram. It seems like short pulsations are difficult to detect with the spectrogram, consistent with the remarks about good time resolution of the scalogram at high frequencies.
Figure 5.26: Spectrogram and scalogram of magnetic fields SC2 2002-08-17.

Filtering is applied to the electric field data to isolate the three pulsations and this is shown in figure 5.27. With the frequency estimators, we obtain three values for the existing frequencies: 5.8, 25.7 and 40.3 mHz with the MUSIC method and 5.8, 25.9 and 40.7 mHz with the eigenvector method. Instantaneous frequencies are obtained for each pulsation with the help of analytic signal. Just as before, there are oscillations around the 5, 16 and 27 mHz components but the oscillations occur a lot more often than in the previous satellite data since we have more pulsations and frequencies involved in this event. A fact worth mentioning here is that not only Pc5 pulsations are present; Pc4 and Pc3 also exist in this event.

Since three different pulsations are detected with the above methods, it would be interesting to perform a wavelet decomposition of the data which allows us to look at the data and its properties in different frequency bands. In this way, we don’t lose any information which could be lost due to choice of bad filters based on the estimation of frequency contents using other methods. The results of wavelet decomposition of level four together with the original and filtered data are plotted in figure 5.30 for comparison. It is easy to see that d4 corresponds to the 6 mHz pulsation which has its maximum and strongest oscillation starting at 11:50. d4 is very similar to the pulsation with approximately 16 mHz frequency. By visual inspection, d3 seems to have 45 cycles
Figure 5.27: Duskward electric field SC2 2003-09-21.

Figure 5.28: The FT and PSD plot of duskward electric field SC2 2003-09-21.

in half an hour interval or about 25 mHz which is consistent with what we see in the scalogram at around 11:30 and 12:50. The pulsation shown in (c) of figure 5.30 is obtained after filtering which looks quite different from d3. It has an artificial look and some small amplitudes seem to be smoothed out and lost. This leads to the conclusion that wavelet decomposition is preferred over direct filtering. One simply choose to look at the data in different frequency bands instead of specifying which frequency interval
one wants to look at in the case of filter operation which can lead to loss of information. The modification of the amplitudes in different levels of the wavelet decomposition seen earlier in the 2nd test signal and probably here as well is not as severe as direct filtering.

5.2.5 Analysis of data from 30th of August, 2001

Finally we examine an event from 30th of August, 2001. Measurements of the duskward electric field for SC2 are shown in figure 5.31 and from the shape of the signal we suspect that it is a sum of two frequency components, especially in the beginning. In the Fourier transform and the PSD plot, we observe some peaks in the interval 5 and 15 mHz with the highest peak at around 6 mHz in figure 5.32. The spectrogram and scalogram are found in figure 5.33. The pulsation with the highest energy occurs at 14:10 at about 7 mHz in both time-frequency graphs. It is more evident in the scalogram compared to the spectrogram that this 7 mHz pulsation occurs a few times more until 14:50.

Looking for pulsations in the higher frequency range, a clear red clump sticks out from 14:10 in the frequency range of 25 and 33 mHz in the two plots of figure 5.32. This is again observed at 14:50 in the scalogram for a very short period. Referring to the FT and PSD plots for some indication of high frequency components in this range, we gain no helpful information. Indeed, one tends to overlook the low amplitude around that frequency range as leakage. Therefore wavelet decomposition is performed hoping for more information about this event. The result is shown in figure 5.34. Here it is obvious that a3 corresponds to the pulsation of frequency around 7 mHz observed by all methods. The d3 corresponds to the pulsation of 27 mHz with their highest amplitude at around 14:10 and 14:50 consistent with observations in the scalogram. Lastly, we
Figure 5.30: Duskward electric field and its wavelet decomposition SC2 2003-09-21. (a) Original satellite data. (b) Bandpass filtered 10-18 mHz. (c) Band-pass filtered 25-33 mHz. (d) Low-pass filtered 8 mHz. $a_4$ denotes approximation at level for and $d_4, d_3,$ and $d_2$ denote the details of the data at level 4, 3, and 2.

refer to the frequency estimation methods which give estimates of 6.4 and 28.6 mHz for MUSIC and the other one 4.9 and 28.3 mHz.
Figure 5.31: Duskward electric field SC2 2001-08-30.

Figure 5.32: The FT and PSD plot of duskward electric field SC2 2001-08-30.
Figure 5.33: Spectrogram and scalogram of duskward electric field SC2 2001-08-30.
Figure 5.34: Duskward electric field and its wavelet decomposition SC2 2001-08-30. (a) Original satellite data. (b) Band-pass filtered 22-32 mHz. (c) Band-pass filtered 25-33 mHz. (d) Low-pass filtered 8 mHz. a3 denotes approximation at level three and d3, d2, and d1 denote the details of the data at level 3, 2 and 1.
Chapter 6

Conclusions and Discussions

The methods presented in this thesis offer us different ways to view the same data and thus have their own advantages and drawbacks. The Fourier transform is a classic in analyzing signals and offers a simple tool for identifying the frequency contents of a signal. It has proven to be quite effective for example for the cases of monochromatic events on the 14th of August, 2003 and 17th of August, 2002. The frequency of those events is roughly constant with a few occasional oscillations. This yields a high distinctive peak in the Fourier transform. The frequency contents of low energy short time pulsations can sometime be overlooked in the FT plots such as the event on the 30th of August 2001. The high frequency pulsation around 27 mHz is visible in the FT but more likely to be interpreted as effects of leakage.

The spectral estimation methods exhibit more or less the same characteristics as the FT since they are based on the Fourier transform. Leakage is reduced to some extent. There is almost no significant difference in terms of performance between the Welch and Thomson multi-taper method on the simulated data as well as the satellite data. The periodogram performs quite well on the simulated data with the sharpest peak but however, on the satellite data, it is outperformed by Welch and Thomson due to leakage.

The parametric methods do not have the same performance as the non-parametric ones in general, especially Burg’s method. As remarked earlier in the first test signal where the frequency is changing very fast and also in the estimation of power spectrum density of the magnetic field on 17th of August, 2002, the Burg estimator tries to represent all frequency components by sharp peaks. The accuracy of the frequency estimation methods partly relies on the fact that the number of the sinusoids is known prior to estimation. The number of sinusoids needs to be estimated based on observations from other methods which can contribute to more uncertainties. It’s very difficult to pinpoint exactly how well and under what circumstances this method performs well. To summarize, we can say that the models offered by the parametric methods can’t sufficiently describe the non-stationary and turbulent nature of the ULF pulsations.
The analytic signal is very sensitive to changes in frequency and thus provides a good tool to visualize the frequency contents of the pulsations in focus. In order to obtain good performance, a quasi-monochromatic pulsation is required. In general, the conventional FT and PSD methods. We can locate the energy in the frequency and evaluate the analytic signal as detected by the short-time FT and PSD plots. When this is done, frequencies and magnetic field can be readily detected in the spectrogram. These methods together with the analytic signal have provided a good tool to follow changes in frequency as function of time and location. The advantage of this method is that in the beginning and especially towards the end of the data, the instantaneous frequency can be subjected to transients and various oscillations.
looking at the pulsations in more depth and details as a function of time. One should avoid using the parametric methods as commented earlier because pulsations with fast changes in frequency and durations can not be modeled efficiently. In this thesis, we have focused on four events from four different days and it would be appropriate to verify the conclusions we draw now by studying more satellite data. A better insight of the physical background of these pulsations and the plasma during the event will also facilitate the analysis of their frequency contents.
Bibliography


