4

Observations of Waves in Space

4.1 Early Field Observations

OBSERVATIONS ON-BOARD rockets and satellites are of fundamental importance for space science. Progress in the area of wave phenomena in the magnetosphere is closely linked to the ability of measuring oscillating quantities in space. Let us spend some time recalling the early achievements in this field.

At the begin of the Space Age, the focus was on getting a general survey of the outer magnetosphere. It was no doubt at the time that observations of both electric and magnetic fields were of vital importance, but the primary objectives of the missions were usually observations of the static, or nearly static, fields. Interesting wave observations were achieved more or less as by-products.

Observing the magnetic fields was the easy part, since measurement techniques were already well developed. So-called fluxgate magnetometers were installed in aircraft and used for submarine detection during World War II [26, 27]. Magnetometers were also carried by several sounding rockets prior to the satellite age [28].

As satellites became available, accurate measurements of the geomagnetic field in the outer parts of the magnetosphere, not reachable with rockets, were eagerly awaited. The first magnetometer flown above 200 km was carried by the Russian satellite Sputnik III, launched in May 1958. The outcome of the experiment was observations of the geomagnetic field up to 800 km, which were consistent with the theoretically predicted field. The objectives of the Pioneer I satellite mission, some months later, were to reach escape velocity and get so close to the moon that scientifically relevant data could be returned. Pioneer did not succeed, but its magnetometer returned data from the outer parts of the magnetosphere that were interpreted as wave motions [27, 29]. The first observations of whistler waves in space were made by Vanguard III as a by-product from the magnetometer used for precise recordings of the background field [28, 30].

To observe the electric fields was a bigger challenge. The fields are weak and the perturbations created by the spacecraft itself and its payload are considerable. Instruments that worked perfectly well in a uniform atmosphere did not function properly in a plasma environment. Although the test flights were unsuccessful the efforts to find new and better measurement techniques continued and two very different ideas were developed parallel in time [31–33]. One of them was a method where a visible plasma cloud was injected into the region of interest. The cloud was then optically tracked and its velocity determined. In a first approximation the electric field could then be obtained from

$$\mathbf{E}_{\perp} = -\mathbf{v}_{\perp} \times \mathbf{B}, \tag{4.1}$$

where $\mathbf{E}_{\perp}$ is the electric field component transverse to the magnetic field, $\mathbf{v}_{\perp}$ is the perpendicular component of the ionized cloud velocity and $\mathbf{B}$ the magnetic field. Test showed that Barium was a suitable element to use and successful rocket flights were performed in 1966-67 [33–36].

The other method was the double probe technique that will be discussed in more detail in the next section. The idea is to measure the potential difference between two points in the plasma [31, 32]. In October 1966 this technique was used to determine electric fields in the auroral ionosphere [37].

As has been pointed out, wave observations were not the primary target for the first space missions, but sometimes a wave instrument was included in the payload. The first satellite experiments originally designed for wave observations were the electric dipole antenna carried on-aboard Aloulette 1 [38] and the magnetic loop antenna on Injun 3 [39]. These instruments were designed for observations of whistlers and other very-low frequency (VLF) radio signals, already recorded by ground based receivers. During these and other missions it was discovered that there were a variety of new wave modes, never previously observed. Since then we have gradually learned that waves play a fundamental role in all magnetospheres and other space plasmas. Today wave instruments are included on virtually every plasma mission [40].

4.2 Techniques for Observing Wave Fields

The techniques for observations of waves are based on very simple principles, although the realization may not be that easy. Variations in the magnetic field are detected using magnetic loop or search coil antennas. They are both based on the law of induction, which states that in a circuit of N turns the voltage ε induced is given by

$$\varepsilon = N\frac{d\Phi_m}{dt}. \tag{4.2}$$

Hence, variations in the magnetic flux Φ_m cause a voltage signal that can be translated into a time series of one component of the magnetic field.[1]

The difference between a loop antenna and a search coil antenna is that the search coil uses a high permeability core to concentrate the magnetic flux trough the circuit. Figure 4.1 provides sketches showing the principles of the two antennas. The loop antenna responds to the magnetic field component perpendicular to the plane of the loop. The search coil antenna responds in a similar way to the magnetic field perpendicular to the plane of the multiple loops. The observed field is then parallel to the axis of the metal rod, as is indicated in Figure 4.1.

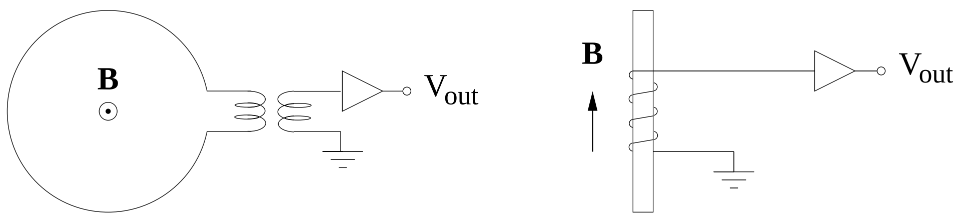

FIGURE 4.1 The principles of a loop antenna (to the left) and a search-coil antenna (to the right). The transformer in the loop antenna design introduces a low frequency cutoff.

One of the challenges when designing a magnetic antenna is to maximize the frequency range for which the instrument can be used. The problem with a loop antenna is that the transformer introduces a low-frequency cutoff that can be as high as 50 Hz. In a search-coil antenna the transformer can be avoided and the low-frequency limit is determined by the basic $d\Phi_m/dt$ response of the coil [40]. As very low frequency modulations of the magnetic field can not be observed by the techniques described, wave observations should include other instruments as well. A flux-gate magnetometer is often flown on the

[1] An inverse technique can be used to change the attitude (orientation) of the satellite in a known magnetic field. If you let a current flow through a coil the Lorentz' force on the moving electrons can actually be strong enough to turn the whole spacecraft.

spacecraft to measure the static magnetic field and such an instrument can be used for observations of low frequency waves also.

To measure the electric wave fields two different designs have evolved: cylindrical dipoles and spherical double probes [40]. They differ in the geometry of the sensing elements, but share the same basic idea. Two conducting elements are placed in the plasma and the potential difference between them is measured. The cylindrical dipole consist of two cylindrical elements extending outward in opposite directions from the spacecraft body. Figure 4.2 shows a sketch of the spherical double probe, where the elements are two conducting spheres connected to the spacecraft with booms.

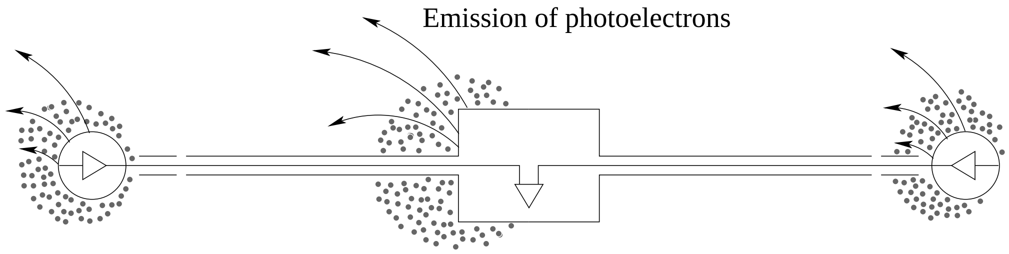

FIGURE 4.2 The principle of a spherical double probe. Two conducting spheres are attached to the satellite with booms. Typical sizes of the spheres and the separation between them are 0.1 m and 100 m respectively. Emission of photoelectrons due to solar UV radiation is indicated. The sun is assumed to be located to the left of instrument.

The electric field in the plasma frame of reference can be found from the relation

$$(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d} = \frac{V_1 - V_2}{|\mathbf{d}|}, \tag{4.3}$$

where $\mathbf{d}$ is the probe separation. The $\mathbf{v} \times \mathbf{B}$ term is the electric field from the motion of the probe system across the Earth's magnetic field [41]. To understand the difficulties involved using this measurement technique, it is important to realize that the probes will generally not have the same potential as the surrounding plasma. When the probe is placed in a plasma it will initially encounter more electrons than ions, as the electrons are more mobile. The net current between the plasma and the probe results in an equilibrium situation (i. e., no net current) where the probe has a negative potential. The measurement still works if the two probes respond identically to the plasma. However, if the spacecraft is sunlit an asymmetry is created. Due to UV radiation both the probes and the spacecraft body will emit photoelectrons. The asymmetry is obvious from Figure 4.2 where the sun is assumed to be located to the left of the instrument. It is easier to observe the electric field in the direction perpendicular to the sunward direction, since the photoelectron clouds then are symmetric with respect to the measurement equipment. In a magnetospheric plasma with low density (< 10 cm^3) the effect of the photoelectrons is large and the instrument is often operated with a bias current from the probe to the spacecraft body. The probe is then brought closer to the plasma potential and reliable observations of almost static and low-frequency waves can be made [42, 43].

Above we have assumed that the waves we want to observe have wavelengths larger than the separation between the probes, $\lambda \gg |\mathbf{d}|$. It should be clear that if the wavelengths are shorter or comparable to $|\mathbf{d}|$ the response of the instruments are much more difficult to interpret correctly [40].

4.3 Spectral Analysis

The output from a wave experiment is usually data in the form of a time series. An example of a time series is given in Figure 4.3. It shows one of the magnetic field components observed by the FAST satellite at 4100 km altitude in 1997 (cp. Paper II). The presented data is despun, that is, the effect of the satellite spin is removed. The

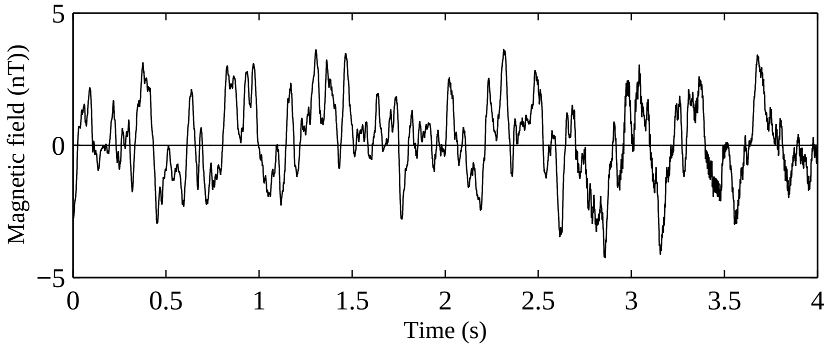

FIGURE 4.3 A typical time series from a wave experiment. This is a magnetic wave field component perpendicular to both the background magnetic field and the satellite velocity. Time is in seconds starting at 1997-04-23/19:47:37.000 UT.

other field components look roughly the same and we realize immediately that analyzing the wave data will be a rather messy task.

To simplify things we recall that almost every function can be expressed as a integral of harmonic waves of different frequencies. There are efficient methods for computing the Fourier transform of a discretely sampled time series. The Fourier transform is in general complex and we still need to reduce the amount of information. Computing $B_x(\omega)B_x^*(\omega)$ and averaging in time we get a real quantity called the power spectral density or simply power (auto) spectrum. Information is reduced in the averaging procedure and the averaged quantities, for example, $< B_x(\omega)B_x^*(\omega) >$ are possible to interpret.

The power spectral density gives a nice overview of the important frequencies hidden in the time series. Sometimes these frequencies can immediately be recognized as some fundamental frequency of the plasma. Figure 4.4 shows an example from [44]. It shows

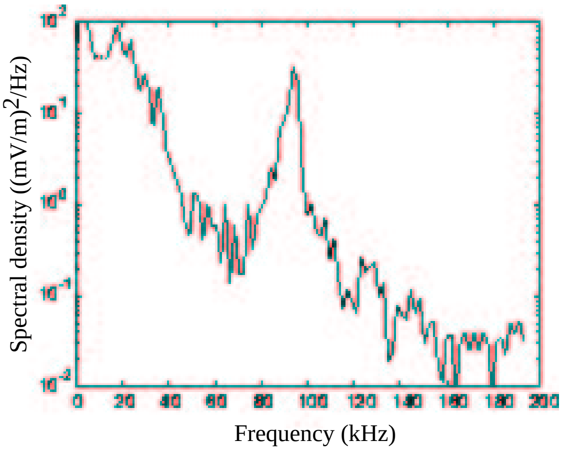

FIGURE 4.4 Electric power spectral density observed by Freja on September 19, 1993, 0916:51.2 UT. The peak at 90 kHz corresponds to the plasma frequency. This gives a density of 100 cm^{-3}.

an electric field power spectrum observed by the Freja satellite. There is a pronounced peak at 90 kHz. With a little knowledge in advance of the densities in the magnetosphere and thus of typical plasma frequencies in different regions, it is possible to conclude that the peak corresponds to the plasma frequency. This information can now be used for a more precise determination of the local density as discussed earlier.

Other characteristic features of in power spectra may reveal more a about the plasma composition. Oscarsson et al. [44] continues the analysis by looking at lower frequencies. The electric field power spectral density for wave fields up to 16 kHz is shown in Figure 4.5. There is a sharp cutoff at just below 2 kHz, which is interpreted as the so-called lower hybrid frequency. From this information the number density ratio between oxygen ions and protons, n_{O+}/n_p can be determined.

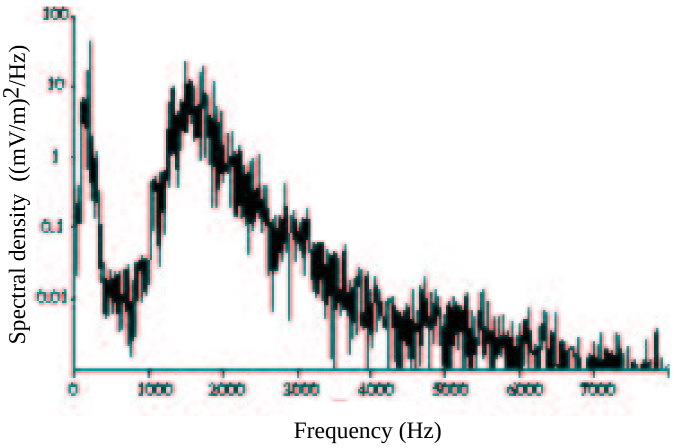

FIGURE 4.5 Power spectrum of the electric field recorded by Freja, showing the low-hybrid cutoff below 2 kHz. The knowledge of the lower hybrid frequency made it possible to determine the ratio n_{O+}/n_p.

To study the polarization of observed waves a natural first attempt is visualize the motion of the wave vector. If the wave is nicely circularly polarized the field vector should rotate in a circle around the ambient magnetic field. To see the wave vector motion, we can plot the component of the wave field perpendicular to the background magnetic field at different times. This type of figure is called a hodogram. An example of a hodogram of a common type is presented in Figure 4.6. We see that it is not always easy to analyze a

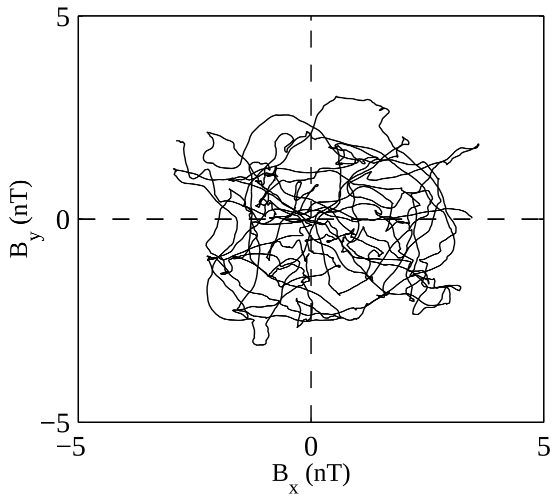

FIGURE 4.6 Hodogram using the first 3000 points of the FAST data shown in Figure 4.3. The wave field vector perpendicular to the magnetic field is plotted. No conclusions concerning the polarization properties should be drawn from this figure.

hodogram. In a space plasma there could be multiple wave modes present simultaneously and a broad spectrum of frequencies. Hence, the wave vector motion could look random. A hodogram that is easier to interpret can be seen in Figure 6.4.

A slightly more advanced way to analyze the polarization is to look at the cross spectra, that is, for example $< B_x(\omega)B_y^*(\omega) >$. The phase difference between x and y component is reflected in the cross spectrum [45]. To see this we write

$$< B_x(\omega)B_y^*(\omega) >= B_R(\omega)\exp(i(\theta_{xy}(\omega)) = B_R(\omega)\exp(i(< \Delta\phi(\omega) >)), \tag{4.4}$$

where $< \Delta\phi >=< \phi_x - \phi_y >$. The phase, $\theta_{xy}(\omega)$ can be computed from

$$\tan\theta_{xy}(\omega) = \frac{\mathrm{Im}[< B_x(\omega)B_y^*(\omega) >]}{\mathrm{Re}[< B_x(\omega)B_y^*(\omega) >]} \tag{4.5}$$

Clearly the cross spectra between two components contain information on the phase difference between them. Sometimes a close investigation of the cross spectrum can give very clear results. An example is given in Figure 6.5, where θ_{xy} is plotted versus ω. The right-hand circular polarization is obvious for waves in the frequency interval 70-150 Hz. Another interesting example is found in [46]. Based on a polarization study of this type it was realized that the search coil magnetometers on the Freja satellite were not mounted as documented.

4.4 Problems with Space Measurements

To observe waves in space is associated with a number of more or less difficult problems. Although many of these really deserve an extensive treatment, we will only briefly comment on some of them in this section.

The first of the problems we will discuss is the problem of making the observations themselves. The huge variety of waves in space pose severe requirements on the experiments used. The frequency range of interest is large, extending from below the ion gyrofrequency to above the plasma frequency. In the Earth's magnetosphere the plasma frequency can be as high as 10 MHz (in the ionosphere) and the proton gyrofrequency as low as 0.1 Hz (in the outer regions). Moreover the intensities of the wave emissions vary a lot and the instrument must be able to detect both high and low intensities [40]. Apart from these scientific requirements the instruments must of course be light-weight, have very low power consumption and be unaffected by all the strains that are connected with space flights [27].

Another inevitable problem is that the spacecraft itself affects the environment that we want to observe. We cannot prevent the plasma from interacting with the spacecraft. One of consequences is that the spacecraft will get electrostatically charged, just as the probes used to measure the electric wave field. Another effect resulting in erroneous observations is the constitution of a wake behind the moving satellite. As the satellite moves through space it sweeps away the plasma and the density will be lower behind the satellite. It is also of vital importance that the satellite is kept electrically and magnetically clean to make sure we are observing the environment and not noise from the electronics of the instruments themselves. There is always the risk of different instruments interfering with each other.

Even if the instruments worked flawlessly and the other effects described can be compensated for, problems still remains. Telemetry is always limited and the amount of data transmitted to the ground station is always less than the actual measuring capacity. We must be careful when deciding which data to transmit, especially if we are interested in wave observations. The reason for this is so-called aliasing. Imaging that we are sampling from a signal, that is a mixtures of different frequencies. For every sampling interval Δt there is critical frequency called the Nyquist frequency given by

$$f_N = \frac{1}{2\Delta t}. \tag{4.6}$$

If the signal we are sampling from contains no frequencies higher then f_N we do not have to worry. However, if higher frequencies are present there is a problem. From a sampled signal it is impossible to distinguish between a wave of the frequency $f_N + \Delta f$ and $f_N - \Delta f$. If we compute a power spectrum all the power spectral density above f_N will be translated as frequencies below f_N. Hence, to reduce telemetry from the spacecraft we cannot just send every second data point. Normally, high frequencies are filtered out on-board the spacecraft, so that highest frequency present in the received data is known. Note that not even the wave instruments record a continuous signal; there is always sampling involved.

The last problem to be discussed is fundamental. We cannot place a satellite at rest with respect to Earth at an arbitrary location in the magnetosphere. To avoid falling down, the satellite must orbit the Earth with a certain velocity. Now, imagine that the instruments detect sudden changes in the plasma as the satellite crosses some interesting boundary in space. How do we determine if the boundary is fixed in space or if it moving with respect to the satellite? How do we know that it has not vanished totally half a minute later? Multi-spacecraft missions such as Cluster (see Chapter 6) try to overcome problems like this.

The wave observations are also more difficult to interpret than we have pretended so far [47]. Normally we interpret the observations as a time variation as long as we do not have any further information. We could observe a standing wave and it would still look like something that oscillates in time. Since the satellite is moving with respect to the plasma, the observed waves will be Doppler shifted. The frequency we observe ω_{sat} is not the frequency of the wave in the plasma frame of reference but related to it through

$$\omega = \omega_{\text{sat}} + \mathbf{v}_{\text{sat}} \cdot \mathbf{k} \tag{4.7}$$

In the atmosphere the Doppler shift is not complicated since for light and sound waves $\omega(\mathbf{k})$ is a known linear function of $\mathbf{k}$. As long as we know the velocity of the observer and the observed frequency we can deduce the frequency in the rest frame. In the plasma, however, the relation between ω and $\mathbf{k}$ is complicated. One frequency can match with many different wave vectors and one wave vector can correspond to different frequencies. Hence, a large number of frequencies and wave vectors in the plasma frame might be consistent with the observations.