Reprinted from *Analysis Methods for Multi-Spacecraft Data*
Götz Paschmann and Patrick W. Daly (Eds.),
ISSI Scientific Report SR-001 (Electronic edition 1.1)

— 1 —

Spectral Analysis

ANDERS I. ERIKSSON
Swedish Institute of Space Physics
Uppsala, Sweden

1.1 Introduction

Large amounts of data, like one or more time series from some spacecraft carried instruments, have to be reduced and presented in understandable quantities. As physical theories and models often are formulated in terms of frequency rather than time, it is often useful to transform the data from the time domain into the frequency domain. The transformation to frequency domain and application of statistics on the result is known as spectral analysis.

The literature on spectral analysis is voluminous. In most cases, it is written by experts in signal processing, which means that there are many texts available outlining the fundamental mathematics and elaborating the fine points. However, this is not the only background needed for a space physicist who is put to the task of actually analysing spacecraft plasma data. Simple questions on normalisation, physical interpretation, and how to actually use the methods in practical situations are sometimes forgotten in spectral analysis texts. This chapter aims at conveying some information of that sort, offering a complement to the textbooks rather than a substitute for them. The discussion is illustrated with idealised examples as well as real satellite data.

In order not to expand the chapter into a book in itself, we concentrate on the application of basic Fourier and wavelet methods, not treating interesting topics in time series analysis like stationarity tests, filtering, correlation functions, and nonlinear analysis methods. Higher order methods like bispectral analysis are also neglected. Fundamentals of such items are covered by many of the references to this chapter, and methods particularly suited for multipoint measurements are of course found throughout this book. Other introductions with multi-spacecraft applications in mind can be found in the paper by *Glassmeier and Motschmann* [1995] and in other papers in the same publication.

The disposition of the chapter is as follows. First, we introduce the basic concepts in Section 1.2, where we also discuss time-frequency methods for the analysis of nonstationary signals. The practical implementation of classical Fourier techniques is treated in Section 1.3, while the implementation of Morlet wavelet analysis is discussed in Section 1.4. In Section 1.5, we turn to the simultaneous analysis of two or more signals by the cross spectrum technique, particularly relevant for the analysis of multipoint measurements. Finally, we touch upon the use of parametric spectral methods in Section 1.6.

1.2 Basic Concepts

1.2.1 Fourier Series

Given any process, represented as a function $u(t)$ of time t, we can only study it for some finite time span $t_0 < t < t_0 + T$. Fourier's theorem states that if the signal is finite and piece-wise continuous, it can be written as

$$u(t) = \sum_{n=-\infty}^{\infty} \tilde{u}[n] \, \exp(-2\pi \mathrm{i} f_n t) \tag{1.1}$$

where

$$f_n = n/T \tag{1.2}$$

and the complex Fourier coefficients $\tilde{u}[n]$ are given by

$$\tilde{u}[n] = \frac{1}{T} \int_{t_0}^{t_0+T} u(t) \exp(2\pi \mathrm{i} f_n t) \, \mathrm{d}t \tag{1.3}$$

The transformation 1.3 replaces a continuous infinity of values $u(t)$ on a finite interval $0 \leq t < T$ by an infinite series of values $\tilde{u}[n]$ for all integers n. Each term in the Fourier sum 1.1 corresponds to an oscillation at a frequency $f_n = n/T$. As such oscillations are the eigenmodes of small-amplitude perturbations in most physical systems, this kind of transformation is an obvious means for understanding the physics behind some signal.

One may note that the coefficients $\tilde{u}[n]$ and $\tilde{u}[-n]$ corresponds to oscillations with the same value of $|f_n|$, which is the quantity determining the time scale of the oscillation. Hence, the sum of the two terms for n and $-n$ in equation 1.1 will describe the part of the signal corresponding to a certain time scale $1/|f_n|$, so for a real signal, this sum must be real. Thus,

$$\tilde{u}[-n] = \tilde{u}^*[n] \tag{1.4}$$

where the star denotes complex conjugation. For a real signal, the two terms labelled by $\pm n$ may be considered to represent the same frequency $|f_n|$. While the $\tilde{u}[n]$ coefficients are sometimes called Fourier amplitudes, the sum of the two $\pm n$ terms in equation 1.1 evaluated using equation 1.4 shows that the amplitude of the sinusoidal wave of frequency $|f_n| > 0$ in equation 1.1 in fact is $2\,|\tilde{u}[n]|$.

1.2.2 Parseval's Relation and Power Spectral Density (PSD)

Central to the physical interpretation of the Fourier series is Parseval's relation

$$\frac{1}{T} \int_{t_0}^{t_0+T} u^2(t) \, \mathrm{d}t = \sum_{n=-\infty}^{\infty} |\tilde{u}[n]|^2 \tag{1.5}$$

which in the case of a real signal, where equation 1.4 applies, becomes

$$\frac{1}{T} \int_{t_0}^{t_0+T} u^2(t) \, \mathrm{d}t = \tilde{u}^2[0] + 2 \sum_{n=1}^{\infty} |\tilde{u}[n]|^2 \tag{1.6}$$

The left-hand side is an average of what we may call the average signal energy or signal power[1]. This nomenclature is not without foundation: for instance, if $u(t)$ is a component of the electric field, magnetic field, or plasma velocity, this signal energy is related to the contribution to the physical energy density (SI unit J/m^3) in the plasma from this field component by a factor of $\varepsilon_0/2$, $1/(2\mu_0)$, or $\rho/2$, respectively, where ρ denotes the mass density. Parseval's relation opens up the possibility of interpreting each of the terms in the sum to the right in equation 1.6 as the contribution to the signal energy from the corresponding frequency f_n.

The terms in the sum at the right of equation 1.6 depend on the length of the time interval T. If this interval is doubled to $2\,T$ by adding the same data set again, i.e. by taking $u(t+T) = u(t)$, $0 \le t < T$, the left-hand side of 1.6 will stay constant, while there will be twice as many terms as before in the right-hand sum, the frequency spacing

$$\Delta f = 1/T \tag{1.7}$$

decreasing by half when T is doubled. On the average, the value of the terms $|\tilde{u}[n]|^2$ in the sum in 1.6 therefore must decrease by a factor of 2. Hence, the coefficients $|\tilde{u}[n]|^2$ depends on signal length, and they are therefore not suitable for describing the physical process itself. To describe the distribution of signal energy density in frequency space, we instead introduce a function S_u known as the *power spectral density (PSD)* by

$$S_u[n] = 2\,T\,|\tilde{u}[n]|^2 \tag{1.8}$$

for all non-negative integers n. Parseval's relation 1.6 then takes the form

$$\frac{1}{T}\int_0^T u^2(t)\,\mathrm{d}t = S_u[0]\,\frac{\Delta f}{2} + \sum_{n=1}^{\infty} S_u[n]\,\Delta f \tag{1.9}$$

Having S_u defined in this way, its value at a particular frequency will not change if changing the record length T in the way outlined above. It is thus possible to picture the discrete set of PSD values as samples $S_u[n] = S_u(f_n)$ of a continuous function $S_u(f)$. In this picture, the Parseval relation 1.9 is the trapezoidal approximation to the integral relation

$$\frac{1}{T}\int_0^T u^2(t)\,\mathrm{d}t = \int_0^{\infty} S_u(f)\,\mathrm{d}f \tag{1.10}$$

Our definition of the PSD has the virtue of having an immediate physical interpretation: $S_u(f_n)\Delta f$ is the contribution to the signal energy from the frequency interval Δf around f_n. It can only be used for real signals, since negative frequencies are ignored.

1.2.3 Phase

As the PSD (equation 1.9) is based on the squared magnitude of the complex Fourier coefficients $\tilde{u}[n]$, it obviously throws away half of the information in the signal. The other half lies in the phase spectrum $\varphi[n]$, defined by

$$\tilde{u}[n] = |\tilde{u}[n]|\ \exp(\mathrm{i}\varphi[n]), \tag{1.11}$$

[1] Signal processing texts sometimes distinguish between the energy and the power of a signal: see the discussion by *Champeney* [1973], chapter 4.

with $\varphi[-n] = -\varphi[n]$ and $\varphi[0] = 0$ due to equation 1.4 The absolute value of the phase depends on exactly when we start sampling the signal, so there is little physical information in one single phase value by itself. On the other hand, there is a lot of information to be gained by studies of the relative phase of two signals, as is further discussed in Section 1.5.

If we know the phase spectrum as well as the PSD, no information has been lost and the signal itself can be reconstructed by using equations 1.11 and 1.8. Constructing time series compatible with some given PSD is sometimes useful in numerical modelling and simulations of space plasma situations. The phase may be selected at random, or by some assumption motivated by the model. One should note that the random phase assumption may result in data with significantly different statistics in the time domain than had the original signal. The physics represented by the two signals can therefore be quite different. Methods for treating such problems exist [see *Theiler et al.*, 1992, and references therein].

1.2.4 Discrete Fourier Transform (DFT)

Our principal interest is in data from instruments providing a sampled rather than continuous output: this replaces the continuous function $u(t)$ above by a discrete set of N measurements

$$u[j] = u(t_j) = u(t_0 + j\,\Delta t) \tag{1.12}$$

where Δt is the sampling spacing, whose inverse is the sampling frequency f_s, and $j = 0, 1, 2, ..., N-1$. From equation 1.2, we get

$$f_n = \frac{n}{T} = \frac{n}{N\,\Delta t} = \frac{n}{N} f_s \tag{1.13}$$

By replacing the integrals in the Section 1.2.1 above by sums, where dt is replaced by Δt, we define the discrete Fourier transform (DFT)

$$\tilde{u}[n] = \frac{1}{N} \sum_{j=0}^{N-1} u[j]\ \exp(2\pi\,\mathrm{i}\,n\,j/N) \tag{1.14}$$

and its inverse

$$u(t_j) = u[j] = \sum_{n=-N/2}^{N/2-1} \tilde{u}[n]\ \exp(-2\pi\,\mathrm{i}\,n\,j/N) \tag{1.15}$$

where we have assumed N to be even (generalisation to odd N is straightforward). The index n here runs from $-N/2$ to $N/2-1$, but it is customary to let it run from 0 to $N-1$ instead, by defining $\tilde{u}[n+N] = \tilde{u}[n]$. This is possible since the only effect of replacing n by $n+N$ in the exponential in 1.15 is a multiplication by $\exp(2\pi\mathrm{i}Nj/N) = 1$. We thus write the inverse DFT as

$$u(t_j) = u[j] = \sum_{n=0}^{N-1} \tilde{u}[n]\ \exp(-2\pi\,\mathrm{i}\,n\,j/N) \tag{1.16}$$

For a real signal, equation 1.4 told us that the negative frequency components carried no extra information, and this is the case for $n \geq N/2$ as well:

$$\tilde{u}[N-n] = \tilde{u}^*[n] \tag{1.17}$$

For the DFT of a real signal, Parseval's relation takes the form

$$\frac{1}{N}\sum_{j=0}^{N-1} u^2[j] = \sum_{n=0}^{N-1} |\tilde{u}[n]|^2 = \frac{1}{2} S_u[0]\frac{f_s}{N} + \sum_{n=1}^{\frac{N}{2}-1} S_u[n]\frac{f_s}{N} + \frac{1}{2} S_u[N/2]\frac{f_s}{N} \tag{1.18}$$

where the PSD estimate is

$$S_u[n] = \frac{2N}{f_s}|\tilde{u}[n]|^2, \quad n = 0, 1, 2, ..., N/2 \tag{1.19}$$

If the signal is measured in some unit X and the frequency unit is hertz (Hz), the unit of the DFT as defined by 1.14 also will be X, while the PSD will be in units of X^2/Hz. This is appropriate, as we may picture the Parseval relation 1.18 as an approximation of the integral relation 1.10.

1.2.5 Normalisation

The definitions 1.14 and 1.16 of the DFT and its inverse are not the only possible. For instance, the choice of in which of the exponentials in equations 1.14 and 1.16 we place the minus sign is quite arbitrary, but for a real time series, this does not change the physical results. The convention adopted here is the familiar one for a physicist, as harmonically oscillating properties normally are represented by functions of the form $\exp(-i\omega t)$ in the physics literature. On the other hand, works in signal processing usually use the opposite sign in the exponentials.

More important is that the factor of $1/N$ we have placed in 1.14 is sometimes placed in front of the sum in 1.16 instead, or even, for greater symmetry, split into two factors of $1/\sqrt{N}$ in front of each of the sums in 1.14 and 1.16. Other factors may also be introduced: see Table 1.1 for a sample of different conventions found in the literature and in some software packages. For a physicist, it is usually natural to think of the Fourier decomposition as a means of separating oscillations at various frequencies, which means that the inverse Fourier transform should be just a sum of the Fourier coefficients multiplied by complex exponentials, as in equation 1.16.

As a further complication, there is little agreement in the literature on how to define the PSD in terms of the Fourier coefficients, i.e., how to write equation 1.19. One reason for this is that many works on spectral analysis are written by mathematicians or experts in signal processing, for whom explicit appearance of sampling frequency in the PSD definition or the confinement to real signals are undesirable. From the point of view of a practising space plasma physicist, a normalisation of the PSD such that Parseval's relation is of the form 1.18 is most often (but not always) the natural choice. As the conventions vary between different algorithms and software packages, it is wise to check the normalisation in the particular routine one is using by straightforward evaluation of both sides of equation 1.18. When publishing spectra, a good service to the reader is to tell what normalisation convention has been used, for example by showing the form of Parseval's relation.

These issues of normalisation may seem dry and uninteresting, but are in fact quite important: somebody will look at the power spectra you publish, try to estimate wave amplitudes in a frequency band or field energy or something similar, and get wrong answers if she/he has misunderstood your normalisation. Trivial as it may seem, the normalisation of spectra is remarkably often a source of confusion.

Table 1.1: Samples of DFT definitions used in the literature and some software packages.

Normalisation conventions in Fourier analysis

Fourier transform: $\tilde{u}[n] = A \sum_{j=0}^{N-1} u[j] \exp(2\pi B\mathrm{i}jn/N)$ Inverse transform: $u[j] = C \sum_{n=0}^{N-1} \tilde{u}[n] \exp(2\pi D\mathrm{i}jn/N)$				
	A	B	C	D
This work	$1/N$	+	1	-
IDL	$1/N$	-	1	+
Mathematica	$1/\sqrt{N}$	+	$1/\sqrt{N}$	-
Matlab	1	-	$1/N$	+
Bendat and Piersol [1971]	1	-	$1/N$	+
Brockwell and Davis [1987]	$1/\sqrt{N}$	-	$1/\sqrt{N}$	+
Brook and Wynne [1988]	$1/N$	-	1	+
Kay [1988]	1	-	$1/N$	+
Marple [1987]	T	-	$1/(TN)$	+
Press et al. [1992]	1	+	$1/N$	-
Strang and Nguyen [1996]	1	-	$1/N$	+
Welch [1967]	$1/N$	-	1	+

1.2.6 Time-Frequency Analysis

In their most straightforward application, the Fourier methods described above transform a function $u(t)$ into a function $\tilde{u}(f)$. Hence, one time series is transformed into one spectrum. This type of transform is sensible if the process one is studying is stationary. However, the space plasma physicist often concentrates on dynamical situations rather than truly stationary phenomena. In addition, the 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. As an example, we may take a spacecraft crossing a magnetopause: the characteristic frequencies and the waves present are different inside and outside the magnetosphere, and in addition there are wave phenomena associated with the magnetopause itself.

In such cases, a more useful representation of the signal will result if we apply the concept of two time scales: one fast time scale, to be represented in the frequency domain, and one slow time scale, for which we keep the time domain representation. We then split the signal record of N samples into M shorter intervals, each of length L, and calculate the PSD for each of these. This technique is known as the short-time Fourier transform (STFT) or the windowed Fourier transform (WFT), as the selection of a particular part of the time series can be seen as a multiplication with a rectangular window function

$$w_k[j] = \begin{cases} 0, & j < kL \\ 1, & kL \le j < (k+1)L \\ 0, & j > (k+1)L \end{cases} \tag{1.20}$$

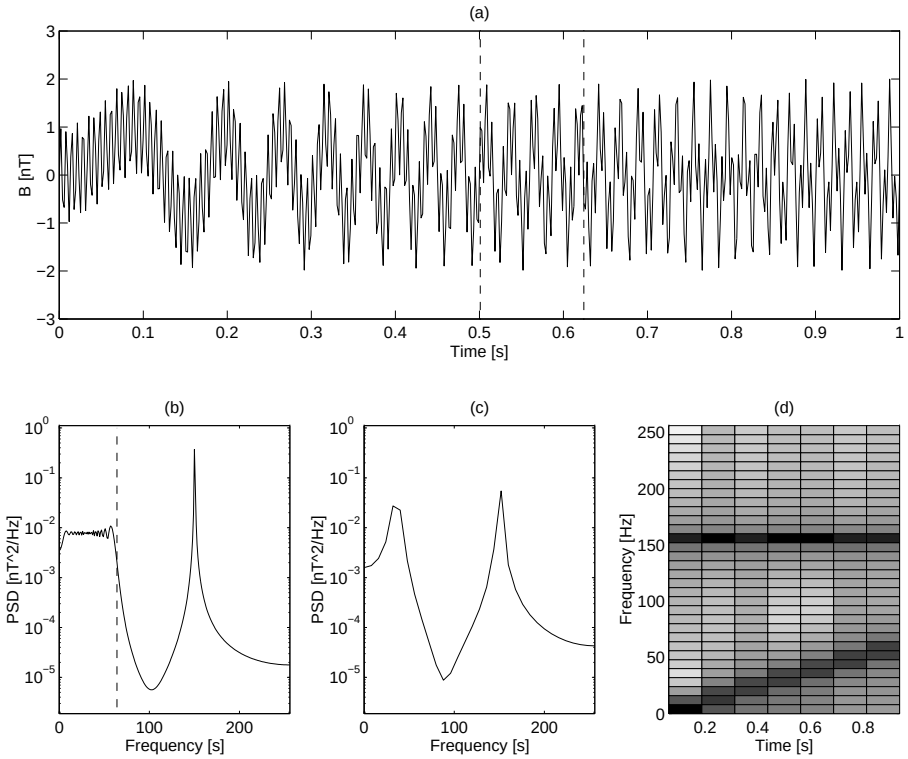

Figure 1.1: Time-frequency analysis. (a) One second of the time series of the signal in equation 1.21, sampled at 512 samples/s. (b) 512-point PSD of the signal in (a). (c) 64-point PSD of the part of the signal between the dashed lines above. (d) 64-point PSDs of 8 subintervals of the signal above. Dark is high and light is low values of log PSD.

Generally, the technique of representing fast variations in the frequency domain and slow variations in the time domain is known as time-frequency analysis. Data analysed in this way are normally presented as colour or grey-scale coded panels, which often are known as dynamic spectrograms. Examples using real data can be found in Figure 1.10 on page 30.

Another example, based on synthetic data, is shown in Figure 1.1. The synthetic signal in panel (a) is represented by $N = 512$ samples of

$$u(t) = B \sin\ 2\pi f(t)t + B \sin\ 2\pi f_0 t \tag{1.21}$$

where $0 < t < 1$ s, $B = 1$ nT, $f_0 = 150$ Hz, and $f(t)$ rises linearly from zero at $t = 0$ to $f_s/8$ at $t = 1$ s. This signal includes a non-stationary part (the first term in equation 1.21) as well as a stationary part (the second term). Panel (b) shows the PSD calculated from equation 1.19. The stationary part is well represented by a narrow spectral peak, while the energy of the non-stationary part is spread out in a frequency band extending up to about $f_s/8$ (the dashed vertical line) without further information. Panels (c) and (d) show

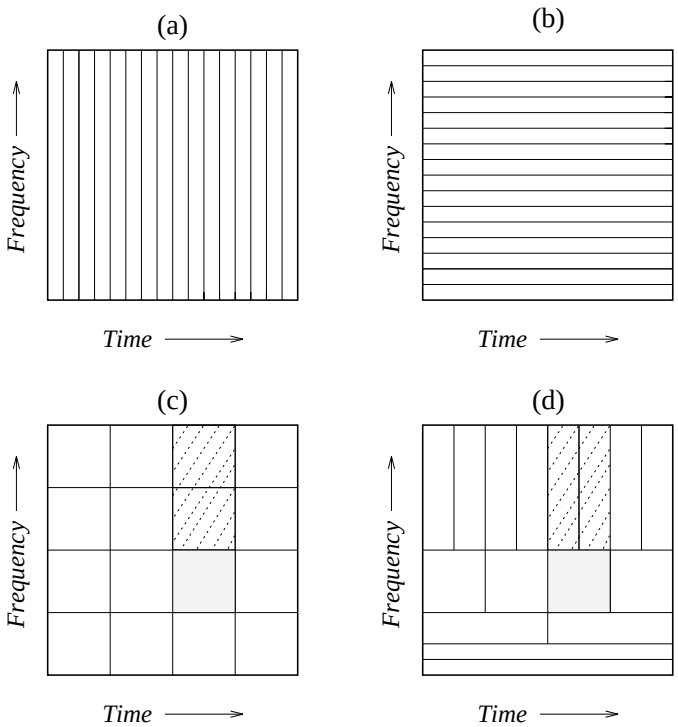

Figure 1.2: Some possible divisions of the time-frequency plane into 16 equal-sized rectangles. (a) Time domain representation. (b) Fourier representation. (c) Time-frequency Fourier representation (WFT). (d) Octave band analysis (wavelet transform). The meaning of the shading is explained in the text on page 28 just after equation 1.43. For an explanation of the dashed rectangles, see page 32, Section 1.4.6.

the effect of dividing the original record into $M = 8$ intervals of length $L = 64$ samples (0.125 s). In (c), we have plotted the PSD (as defined by 1.19) based on the $L = 64$ points between the dashed lines in panel (a). Finally, in (d) we see a time-frequency plot, where the PSD is represented by a logarithmic grey scale.

Obviously, the time-frequency analysis has several merits. In panel (d) of Figure 1.1 we can clearly see the linear trend in the frequency of the low frequency signal, and the stationary character of the high-frequency part. Also, in panel (c), the low frequency signal is represented by a peak with a maximum value of the same order as the peak value for the stationary high-frequency signal, which is reasonable. The low-frequency peak is of course broader, as the frequency of this signal varies from 32 Hz to 40 Hz during the analysed time interval. On the other hand, the representation of the stationary signal at 150 Hz is better in (b) than in (c), in the sense that its energy is more concentrated in a narrow spectral peak.

This illustrates the basic trade-off situation in time-frequency analysis: we can get better resolution in time only at the expense of resolution in frequency. Having a time series of N points sampled at frequency f_s, we get a frequency resolution $\Delta f = f_s/N$ if we calculate the Fourier transform of the full record. In that case, the time resolution is

equal to the length of the record, $\Delta t = N/f_s$, so that $\Delta f\ \Delta t = 1$. Dividing into M records increases the time resolution by a factor of M, while the frequency resolution goes down by the same factor, so that we still have $\Delta f\ \Delta t = 1$. In general,

$$\Delta f\ \Delta t \geq 1 \tag{1.22}$$

where the inequality results when averaging is introduced (Section 1.3.4). As our definitions of Δf and Δt are not in terms of standard deviation as in quantum mechanics [*Landau and Lifshitz*, 1977, p. 47], one should not exactly identify 1.22 with the Heisenberg uncertainty relation, although the two are related.

The relation 1.22 may be illustrated as in Figure 1.2, where we divide the time-frequency plane into 16 rectangles of unit area in four different ways. Panel (a) represents the untransformed time series,[2] and panel (b) the DFT of the full signal. Panel (c) is an example of a time-frequency Fourier analysis.

We could also think of less symmetric ways of dividing the time-frequency plane of Figure 1.2. Panel (d) shows a particularly useful example, known as octave band analysis. Here, the time resolution is different for different frequencies, giving better resolution in time for higher frequency components. This decreases the frequency resolution at high frequencies, but the relative resolution $\Delta f/f$ is kept constant. We could express this as having constant resolution in $\log f$. Implementing this kind of division of the time-frequency plane leads to what is known as wavelet analysis, to which we return in Section 1.4.

Which division of the time-frequency plane in Figure 1.2 is best? The answer clearly depends on the properties of the signal, and the questions we put to it. If the signal, or at least the component of it we are interested in studying, can be assumed to result from a truly stationary process, the resolution is better placed in the frequency than in time, as in panel (b). If, in addition, the signal can be assumed to be composed of a set of discrete frequencies, there may even be reason to consider parametric spectral models, which we briefly touch upon in Section 1.6. A Fourier implementation of time-frequency analysis (panel (c)) is useful for a situation with one slow time scale, represented in the time domain, modulating fast oscillations, represented in the frequency domain. An example of a situation where such a description is useful is a magnetospheric satellite flying through a density cavity, where the slow time scale is set by the size of the cavity divided by the spacecraft speed, while the waves present the fast variations. Finally, the wavelet division of the time-frequency plane in panel (d) is well suited for turbulent situations, with a more or less continuous hierarchy of time scales.

Wavelet analysis can be extended to include other unsymmetric partitions of the time-frequency plane than the one in Figure 1.2d. We touch briefly upon this in Section 1.4.6 (page 32).

As well as depending on the signal properties, the method to use is to some extent dependent on our interpretation of the signal, which in turn may depend on other data, our experience, or our prejudice. Consider the time series of synthetic data in panel (a) of Figure 1.3. The time-frequency analysis in panel (b) suggests a description in terms of a modulated wave at 300 Hz. On the other hand, the single Fourier spectrum of all points in panels (c) and (d) suggests a description as a superposition of two sinusoidal waves at 297 Hz and 303 Hz. Mathematically, these descriptions are of course equivalent. However,

[2] Formally, panel (a) should be interpreted as $M = 16$ DFTs, each based on $L = 2$ data points, from a total record of $N = 32$ samples, since the N-point PSD only gives information at $N/2$ frequencies.

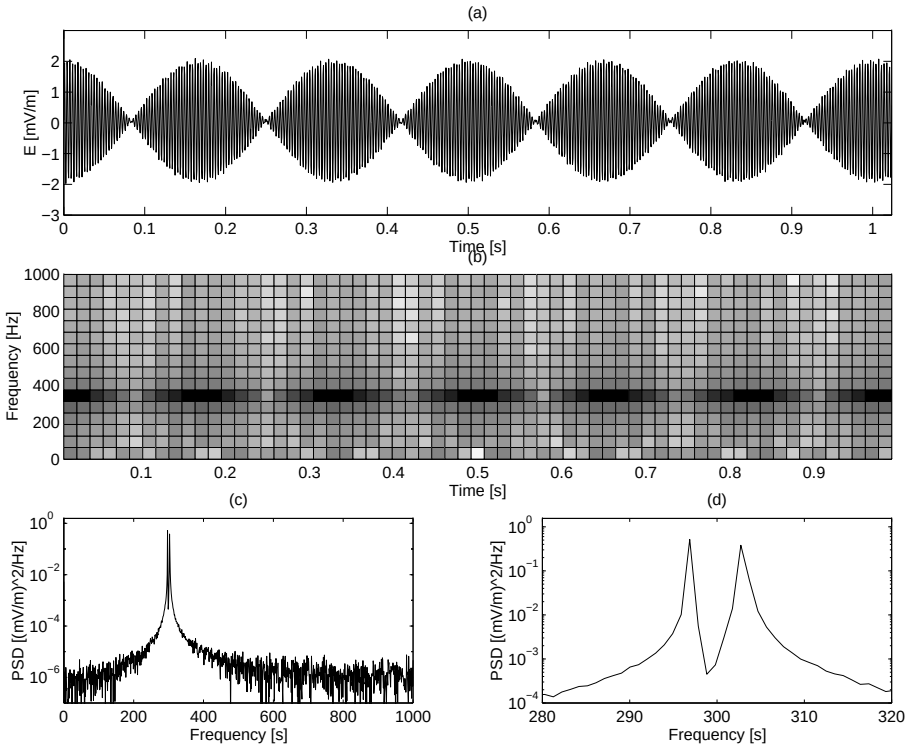

Figure 1.3: A modulated sine wave (a), with some spectral representations: (b) Fourier time-frequency analysis, (c) Fourier spectrum, (d) detail of Fourier spectrum.

one or the other may be more physically meaningful in a given situation. For example, if some other measurement shows that a likely source of the wave is an electron beam whose energy flux pulsates with a frequency corresponding to the observed modulation frequency, a description in terms of modulations, as in panel (b), is appropriate. On the other hand, if we have reason to believe in the presence of two distinct localised and stable sources with slightly different characteristics, the description offered by panels (c) and (d) is more natural.

In the 1990s, wavelet methods (Section 1.4) have became popular. Wavelets constitute a major advance in the spectral analysis, opening new approaches to signal handling, but there is no reason to throw the traditional Fourier based methods overboard. As outlined above, and as will further be discussed in Section 1.4.2, there are applications where a division of the time-frequency plane as in panel (c) of Figure 1.2 is more appropriate than the wavelet division in panel (d). In other circumstances, the wavelet division is the more advantageous.

1.3 Fourier Techniques

This section considers how to practically use Fourier methods for PSD estimation. Ideas and concepts are described, while detailed algorithms are left for textbooks like *Jenkins and Watts* [1968], *Bendat and Piersol* [1971], and *Kay* [1988]. For the reader interested in clear and concise treatments on a level between this section and the textbooks, the old but concise paper by *Welch* [1967] and the practically oriented text by *Press et al.* [1992] are recommended.

1.3.1 Fast Fourier Transform (FFT)

The fast Fourier transform (FFT) is an algorithm for efficient implementation of the DFT 1.14 and its inverse 1.16. This algorithm is the only implementation of the DFT that you are ever likely to see or use, as it is far more efficient than for example a direct summation of the series in equation 1.14. The FFT algorithm is described in most textbooks on signal processing [e.g. *Bendat and Piersol*, 1971] or mathematical computer methods [e.g. *Press et al.*, 1992], and there you will find out that while a direct implementation of equation 1.14 requires a number of calculations proportional to N^2, the clever FFT algorithm only needs on the order of N log N calculations for doing the same job.

The FFT algorithm is included in virtually all software packages intended for data analysis, like Matlab or IDL, so you should never have to actually write the code yourself or even understand it in detail. However, one should be aware that the computational efficiency of the FFT is fully exploited only when the number of data points is an integer power of two ($N = 2^M$).

1.3.2 Detrending

The finite time series from a measurement is often to be regarded as a finite observation of a process, stationary or non-stationary, that was going on before the start of our observations and that will continue afterwards. Hence, it is very likely that our data contains components with periods longer than the length of the data record. Such components in the signal are called *trends*. As the trends represents time scales longer than those a spectral analysis of our data record can resolve, it is desirable to remove them from the data prior to spectral estimation. If not removed, trends may distort the low-frequency end of the estimated spectrum.

Detrending may often be done by fitting (by least-squares methods or otherwise) a linear or perhaps quadratic function to the data, and thereafter subtract this function. Sometimes, more advanced fits like sinusoidals may be called for.

An example of a signal with trends is shown in Figure 1.4. Panel (a) shows 3 seconds of measurements of a spin-plane component of the electric field by a magnetospheric satellite spinning around its axis with a period of 6 seconds. One trend is obvious: there is a sinusoidal variation at the spin frequency, of which we see half a period. This signal is due to spin modulation of the electric field induced by the spacecraft motion through the geomagnetic field, and is indicated by the dashed line. Removing this, we get a time series as in panel (b). Considering the part of the time series between the vertical lines in panel (b), we find another trend as well, due to a field variation at a time-scale longer than the length of this interval. Panel (c) shows an enlargement of this part of the total time series,

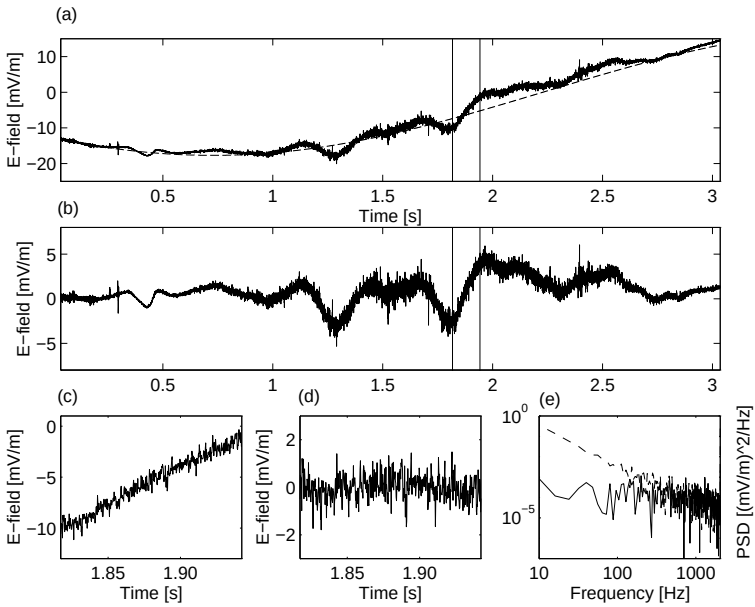

Figure 1.4: Measurement of an electric field component by the wave instrument on the Freja satellite. Zero time corresponds to UT 005416, April 12, 1994. The signal has been low-pass filtered at 1.3 kHz before sampling at 4000 samples/s.

before any of the two trends have been removed. Panel (d) shows the same signal after removal of a linear trend by subtracting a least-squares fit. Finally, panel (e) displays the power spectra of the original signal from panel (c) (dashed) and of the detrended signal from panel (d) (solid), calculated according to equation 1.19 with a 512 point DFT. It is clear that the linear trend carries energies to all frequencies, thereby drowning details in the signal we really are interested in. For example, the peak in the solid curve around 30 Hz cannot be discerned in the dashed curve, where it is completely drowned by the spectral energy from the linear trend.

Removing a trend is a major manipulation of the data. Hence, caution is needed, so that the detrending does not do more harm than good. A few non-physical outliers in the data could wreak havoc in a fitting algorithm, and thus cause the detrending to actually add a spurious component in the data. Before removing a trend, its physical origin should be understood. Visual inspection of the time series before and after detrending is a very good idea.

1.3.3 Windowing

In Section 1.2.6, we introduced the rectangular window function 1.20. The abrupt behaviour of the rectangular window at the edges causes it to leak quite a lot of energy from any main spectral peak to other frequencies. This can be understood by considering that when applying a data window to the time series, we will analyse the signal $w_k[j]\,u[j]$

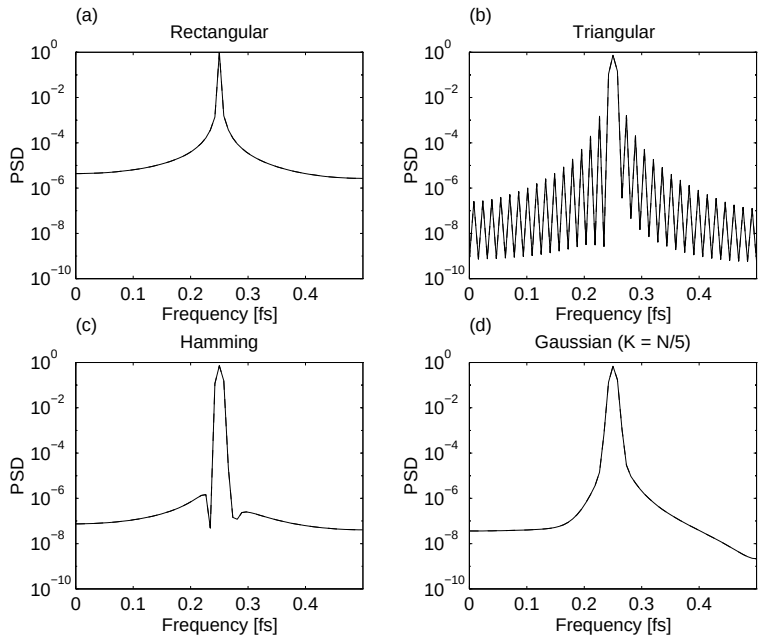

Figure 1.5: 128 point PSDs of a sinusoid at $f = f_s/5$ calculated using different data windows. The frequency is in units of the sampling frequency f_s. The PSD values are normalised so that the total integrated power for $0 < f < f_s/2$ is unity.

rather than the presumably infinite record of $u[j]$ itself. As a multiplication in the time domain is a convolution in the frequency domain, the effect of applying a window will be to convolve the "true" PSD with the PSD of the window. As the Fourier transform of a step function is spread out over all frequencies, this implies that part of the signal energy is moved from any frequency to all others. We illustrate this by an idealised example in Figure 1.5a. The 128 point PSD of a signal consisting of only a pure sinusoid at $f = f_s/5$ is calculated (using a rectangular window 1.20) and normalised so that the total integrated PSD in $0 < f < f_s/2$ is unity. It can be seen that signal energy leaks to the entire frequency interval. The detailed behaviour depends on the frequency of the signal: in the extreme example of an integer number of wave periods fitting into the data interval, there is no leakage at all, as the signal frequency exactly equals one of the Fourier frequencies 1.13 in this case. For a real signal, this is of course a zero-probability event, and a leakage as in Figure 1.5a will generally occur.

As a remedy for the frequency leakage behaviour, a variety of other window functions have been designed, a few of which are shown in Figure 1.6. As can be seen from the idealised example in Figure 1.5, these windows all decrease the leakage of energy to frequencies far away, at the cost of increased width of the main peak. That this price has to be paid is clear from the uncertainty relation 1.22: the window concentrates the signal in time, and hence must gives a widened spectral peak.

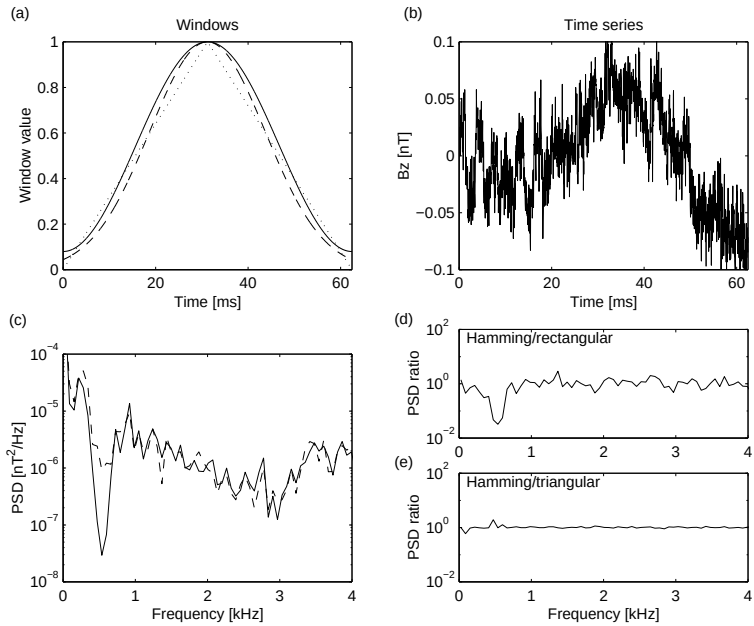

Figure 1.6: (a) Some often used data windows: triangular (dotted), Hamming (solid) and Gaussian ($K = L/5$, dashed). (b) Time series of magnetic field data from the Freja wave instrument. (c) PSD calculations using Hamming (solid curve) and rectangular (dashed) windows. (d) and (e) Ratios of PSDs calculated with different windows.

The simplest of these windows is the triangular (or Bartlett) window, given by

$$w[j] = 1 - \left| 1 - \frac{2\,j}{L-1} \right| \tag{1.23}$$

inside the interval of interest (which is the same as in equation 1.20 above) and zero outside. As is seen in Figure 1.5(b), it has a considerably better behaviour than the rectangular window in terms of energy spread. Even better in this respect and very widely used are the Hamming window

$$w[j] = 0.54 - 0.46\,\cos\left(\frac{2\pi j}{L-1}\right) \tag{1.24}$$

and the Hann (often called Hanning) window (not shown)

$$w[j] = 0.5\ \ - 0.5\,\cos\left(\frac{2\pi j}{L-1}\right) \tag{1.25}$$

Both are zero outside the L-point interval of interest. The Gaussian window

$$w[j] = \exp\left(-\frac{(j - \frac{L-1}{2})^2}{2\,K^2}\right) \tag{1.26}$$

does not go exactly to zero, but may in practice be cut off at some distance from the centre. For an infinite time series, this window is optimal in terms of low leakage to frequencies, which makes it theoretically favoured [*Rönnmark*, 1990]. It differs from the others in that its width K is an adjustable parameter (chosen to $L/5$ in our examples). We will return to the Gaussian window when discussing Morlet wavelet analysis in Section 1.4.1.

The detailed properties of these and other windows have been studied by for example *Harris* [1978]. However, for most applications to data sampled in space, there is little point in putting much effort into the choice of data windows. In practice, the difference between spectra obtained by use of Hamming, Hann and Gaussian windows will be small: even the triangular window is tolerable for many purposes. This is illustrated by an analysis of magnetic field data from the Freja satellite shown in Figure 1.6(b). The PSDs of this time series (using averaging of 4 spectra, see Section 1.3.4) calculated with rectangular and Hamming windows are shown in panel (c). It can be seen that the Hamming window can resolve the minimum near 0.5 kHz much better than can the rectangular window. This difference is illustrated in panel (d), where the ratio of the PSDs in panel (c) is plotted. Finally, in panel (e) we plot the corresponding ratio for the Hamming and triangular windows, and find a much smaller difference. The important point is to use some reasonable window other than the rectangular, and preferably not the triangular.

In general, when multiplying the signal by window coefficients, which all are ≤ 1, signal energy is lost. The exact amount of energy lost will depend on the signal: obviously, a record consisting of a single narrow spike in the middle of the interval, surrounded by zeroes, will lose no energy at all as the window functions all have the value one at the centre, while a record consisting of a spike at one of the end points and all other points being zero will be almost completely lost. By considering Parseval's relation 1.19, we find that the statistically expected decrease in the PSD value due to windowing should be the mean square value of the window,

$$W_{\mathrm{ss}} = \frac{1}{N} \sum_{j=0}^{N-1} w[j]^2 \tag{1.27}$$

In order to compensate the PSD for the energy loss in the windowing, it should be divided by the W_{ss} value for the window at hand.

1.3.4 Averaging and Stationarity

In most texts on signal analysis [e.g. *Kay*, 1988; *Welch*, 1967], it is shown that when applying 1.19 with rectangular window to a time series resulting from a stationary random process, we get standard deviations in the PSD estimate equal to the PSD values themselves, so that the expected error is 100%. This situation is improved (and complicated) by the use of data windows other than the rectangular window (see, for instance, *Welch* [1967], *Jenkins and Watts* [1968], or *Brockwell and Davis* [1987]), but is still not satisfactory.

We emphasise the assumptions made above: (1) the time series is one particular realisation of a random process whose parameters we want to estimate, and (2) these parameters are constant or slowly varying, so that the process is almost stationary. It is only in these circumstances the question of standard deviation enters the problem. If we interpret our data as a unique observation of a deterministic non-stationary process whose details we

want to explore, there is no problem in the fact that consecutive spectra may show large variation—it simply reflects the changing physical situation. We will return to this point below, after briefly having considered the stationary case.

If we actually have an almost stationary random process, our first attempt to decrease the variance in the PSD estimates may be to increase the record length L. However, the effect of increasing the number of points in the DFT is only to get a PSD evaluation at more frequencies, not to improve the estimate at any single frequency, so this approach only gives us more points with the same bad variance. Increasing the record length is still wise, but we should not put all the samples into the DFT at once. Instead, we divide the total record of L samples into P shorter pieces of data, each of length K. After detrending and windowing each K-point interval, we compute its K-point PSD. Finally, we take the average of the P PSDs as our spectral estimate. Averaging reduces the variance with a factor of $1/P$, so in the case of rectangular windows, the normalised standard error will go down from the 100% mentioned above to

$$\epsilon = 1/\sqrt{P} \tag{1.28}$$

Some data analysis packets include means for calculating the confidence interval of the PSD estimate of a given time series, using methods described by e.g. *Jenkins and Watts* [1968]. If such means are not utilised, one could use the rule of thumb that spectral features below the $1/\sqrt{P}$ level should not be trusted to result from a stationary process.

Averaging can also be performed in the frequency domain rather than in time [e.g. *Bendat and Piersol*, 1971]. In that case, one calculates the PSD of all L data points, and then replace the $L/2$ PSD values one gets by $L/(2P)$ values, each being the average of P neighbouring frequency components in the original PSD. The effect on the standard deviation of the PSD estimate is similar to the averaging of spectra in the time domain.

By reducing the amount of information in the original signal by a factor of $1/P$, we increase the quality and comprehensibility of the remaining data. The loss of information in the averaging is described by the uncertainty relation 1.22, which for averaged spectra becomes

$$\Delta f \, \Delta t \geq P \tag{1.29}$$

Averaging is useful for PSD estimates, but not for estimates of the complex DFT 1.19 or the phase spectrum 1.11, as the absolute value of the phase angle is completely dependent on exactly when the data interval starts.

While averaging is useful for reducing noise from statistical fluctuations in a signal from a random process, it is not always justified for a deterministic signal. The idea of improving the PSD calculation is linked to the concept of our signal as a sample signal from a random process. In this view, the PSD is a statistical property of the signal, and we want to estimate it with as little noise as possible. This is often a very reasonable way of looking at the signal, but is not the only possible. For averaging to be physically justified, the signal must be assumed stationary over the time interval from which we construct the PSDs which we average. However, we are sometimes interested in the spectral properties of a non-stationary signal. For example, some sort of wave packet may be observed when passing a spacecraft. There is only this wave packet, so there is no ensemble to average over. Still, its spectral properties can be very interesting. It is more natural to look at this wave packet as a deterministic signal whose physical properties we can investigate in detail than to consider it a sample from a random process whose statistical parameters

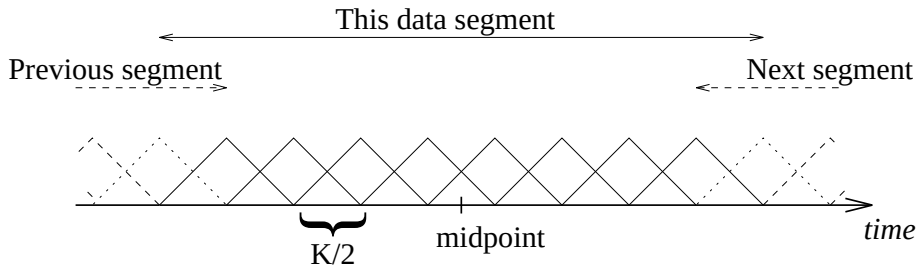

Figure 1.7: Example of use of overlapping time intervals for averaging of spectra. Eight overlapping triangular windows (solid triangles) of length K are applied to the data, the PSD of each corresponding time interval is calculated, and the average is constructed. This is then taken as the power spectrum for the data period denoted "This data segment", which is time tagged by its midpoint. We then proceed to the next segment, whose first triangular windows are shown dashed.

we want to estimate. With this view, "irregularities" in the spectrum are not regarded as "noise" but as physical signals to be explained. A good discussion about conscious and tacit assumptions on stationarity in spectral analysis of space plasma waves is given by *Rönnmark* [1990].

For the analysis of deterministic data with a strong component of high frequency random noise superimposed, averaging may be useful. If the deterministic signal is stationary under the length of the analysis, averaging can reduce the effect of added noise.

Most textbooks on spectral analysis take the statistical, random process approach [e.g. *Bendat and Piersol*, 1971; *Brockwell and Davis*, 1987; *Kay*, 1988]. A basic assumption in these works is that the signal is a sample of a stationary random process. In contrast, stationarity is not emphasised in modern texts on wavelet analysis. This is partly motivated by the ability of wavelets, localised in time as well as frequency, to model non-stationary phenomena (see Section 1.4).

1.3.5 Overlapping Intervals

In the discussion of averaging in Section 1.3.4 above, we took a time series of L points, divided it into segments of K points each, and applied a data window to each segment. Windowing implies giving the data points that are multiplied by the flanks of the data window a low statistical weight in the final result. It is therefore reasonable to use overlapping windows, as illustrated in Figure 1.7. An overlap of 50% is a natural choice, particularly for the triangular window. Except for the first data points of the first segment and the last points of the last segment, each data point will be used in two data windows, and hence in two PSD estimates before averaging. If the weighting of the point is w_1 in one of the windows, it will in the case of triangular windows be $1 - w_1$ in the other window, giving a total statistical weight of 1 for all data points.

In the case of triangular windows and 50% overlaps, the normalised standard deviation in the PSD will not be $1/P$ as suggested by equation 1.28, but [*Welch*, 1967]

$$\epsilon' \approx 1.2/P \tag{1.30}$$

The other commonly used windows 1.24–1.26 give ϵ values of the same order for 50% overlap. The factor of 1.2 may seem discouraging, but one should note that the P in equation 1.30 is not the same as in 1.28, since overlapping makes it possible to squeeze in more windows in a certain amount of time. For a given record which is separated into P non-overlapping subintervals of length K, we can squeeze in $P' = 2P - 1$ windows of length K if we allow a 50% overlap. Hence, P in 1.30 should be replaced by $2P - 1$ if we want to compare the performance of 50% overlapping triangular and non-overlapping rectangular windows on the same data record. Already for $P = 2$, giving $P' = 3$, we find that the ratio ϵ'/ϵ goes down to 0.8. For $P = 5$ ($P' = 9$), this ratio is 0.67. Hence, the use of overlapping windowed data increases the quality of the output in some sense, at the expense of more computation.

One may note that overlapping will never do any harm except for increasing the computation time. In signal processing literature, there is sometimes skepticism against the use of overlapping, which is justified if you want to construct computationally efficient routines. For the practising physicist, who is eager to extract as good information as possible from his/her data, it is often advisable to sacrifice some computational efficiency for optimal use of the data.

We have here discussed overlapping of intervals for which the PSDs are averaged. Another possibility in WFT analysis is to allow overlap of two consecutive data intervals for which averaged PSDs are calculated. By letting this overlap be almost complete, i.e. just skipping one data point between total PSD estimates, one can give the time-frequency analysis an apparent time resolution similar to the time resolution of the original time series. This gives a smoother appearance of a time-frequency plot, which would remove the square pattern from a display like Figure 1.3(b). No effective time resolution is gained, as this is governed by the uncertainty relation 1.29. If using this smoothing method, one should be aware that there may be different weighting of different data points unless the overlapping is done cleverly. We can see this from Figure 1.7, where this kind of smoothing would correspond to letting "previous segment" and the "next segment" move in toward the midpoint. This could result in some data points being used in two PSD calculations, others in three, and others in four PSD calculations.

1.3.6 Zero-Padding

Let us assume that we have a time series record of N samples, which we increase to length $2N$ by adding N zeroes. What happens to the Fourier transform? From the DFT definition 1.14, it is obvious that we get twice as many Fourier coefficients $\tilde{u}[n]$ as before. Also, it can be seen that coefficient number n in the length N expansion must be equal to coefficient number $2n$ in the new length $2N$ expansion, as $u[j]$ is zero for $j > N$. This illustrates that zero-padding, which is the process of lengthening a time series by adding zeroes at its end, makes no essential change to the shape of a spectrum [*Marple*, 1987]. At first, it may seem that zero-padding increases the frequency resolution, as the number of Fourier coefficients increase. However, the amount of information on the real physical signal is obviously the same with or without zero-padding, so the increased frequency resolution must be spurious. For the data from a given time interval, the frequency resolution will always be limited by equation 1.22.

An illustration of the spurious increase in frequency resolution is given in Figure 1.8. The power spectrum of a test signal consisting of two equal-amplitude sinusoids at fre-

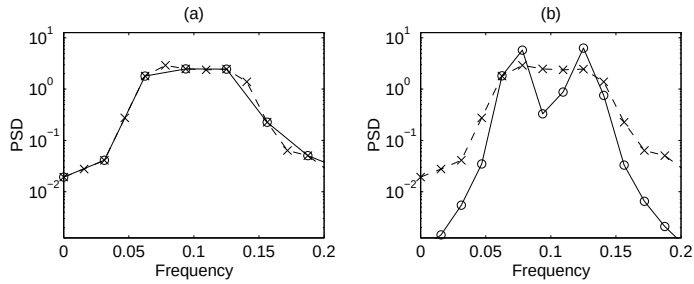

Figure 1.8: Effect of zero-padding. PSDs of a test signal consisting of two sinusoids at frequencies 0.075 and 0.125, in units of the sampling frequency. Dashed lines with crosses (in both panels): 32 data points, 32 zeroes padded. Solid lines with circles (in both panels): no zero-padding, (a) 32 data points, (b) 64 data points.

quencies 0.075 and 0.125 in units of the sampling frequency is calculated from 32 points without zero-padding (solid line in (a)), from 64 points without zero-padding (solid curve in (b)), and from 32 data points with 32 zeroes padded (dashed line in both panels). In (a), we can see that the PSDs of the signals with and without zero-padding coincides at the frequencies where both are evaluated. In (b), the PSD of the zero-padded signal is evaluated at the same frequencies as the 64-point signal, but it cannot resolve the two spectral peaks.

In general, if we pad a signal record of N data points with M zeroes, the Fourier coefficients of the resulting record are given by substituting $M+N$ for N in equation 1.14, where $u[n] = 0$ for $N < j < N + M$. Parseval's relation then takes the form

$$\frac{1}{N+M}\sum_{j=0}^{N-1} u^2[j] = \sum_{n=0}^{N+M-1} |\tilde{u}[n]|^2 \tag{1.31}$$

However, we do not want the padded zeroes to change the physics of the situation as expressed by the PSD estimate. By considering Parseval's relation, we find that the PSD expression 1.19 should be multiplied by a factor $(N+M)^2/M^2$, so that

$$S_u[n] = 2\frac{(N+M)^2}{N}|\tilde{u}[n]|^2/f_s \tag{1.32}$$

is the proper expression for the PSD from a zero-padded signal, evaluated at frequencies

$$f_n = \frac{n}{N+M} f_s, \qquad n = 0, 1, \ldots, N+M-1 \tag{1.33}$$

If combined with windowing, zero-padding should obviously be applied after the windowing. Otherwise, we would not get the smoothness at the ends of the non-zero part of the time series which is one of the aims with windowing. Also, the correction factor 1.27 for the power loss in the windowing process would be erroneous if the non-zero data occupied only part of the extent of the window. When combined with extra overlapping discussed in Section 1.3.5 above, it results in smooth time-frequency spectra without the sharply defined squares of Figure 1.3.

Since zero-padding does not provide any new information, just a smoothing of the spectrum, it is not very useful in most cases. However, in cases where the length of the time series to be transformed by an FFT algorithm is not an integer power of two, zero-padding up to the next such number may actually speed up the FFT calculation.

Zero-padding is not the only way of "cheating" by evaluating the signal at more frequencies than the data really allow. Another variant is by what may possibly be called "continuous Fourier transform", where one evaluates the DFT defined by 1.14 at arbitrary real values of $n < N/2$, not just integers. This scheme is rarely used in Fourier analysis, although it is common in wavelet applications (Section 1.4.3).

1.3.7 Welch Method for Time-Dependent PSD Estimation

Applying the techniques above, we get an algorithm for estimation of the time-dependent PSD. Having all the building blocks at hand, we summarise the resulting method below. This technique is due to and very well described by *Welch* [1967].

1. Divide the total time series

$$u[j], \qquad j = 0, 1, \ldots, N-1$$

of N samples into M shorter intervals

$$u_m[j], \qquad j = p, p+1, \ldots, p+L-1, \qquad m = 0, 1, \ldots, M-1$$

of L samples each. The time resolution of the time-frequency analysis will be L/f_s. If there is an overlap of r points, we have $p = m\,L - r$.

2. Divide each of the intervals of length L into P segments

$$u_{ml}[j], \qquad j = q, q+1, \ldots, q+K-1, \qquad l = 0, 1, \ldots, P-1$$

of length K. If these intervals overlap by s points, $q = p + l\,(K-s)$. A good choice is $s = K/2$.

3. Multiply each data segment term by term by a window function $w[j-q], \quad j = q, q+1, \ldots, q+K-1$.

4. Calculate the DFT

$$\tilde{u}_{\mathrm{ml}}[n] = \frac{1}{N} \sum_{j=q}^{q+K+Z-1} w[j-q]\,u[j]\,\exp(2\pi \mathrm{i} n(j-q)/K), \qquad n = 0, 1, \ldots, K/2,$$

of the windowed time series, preferably using the FFT algorithm.

5. Calculate the PSD estimate, corrected for the windowing:

$$S_{\mathrm{ml}}[n] = \frac{2K}{f_s W_{ss}} |\tilde{u}_{\mathrm{ml}}[n]|^2$$

6. Average over the P short segments to get a PSD estimate with better variance:

$$S_{\mathrm{m}}[n] = \frac{2KP}{f_{\mathrm{s}} W_{\mathrm{ss}}} \sum_{l=0}^{P-1} |\tilde{u}_{\mathrm{ml}}[n]|^2$$

7. $S_{\mathrm{m}}[n]$ is our resulting time dependent spectrum, evaluated at frequencies

$$f_n = \frac{n}{K} f_{\mathrm{s}}$$

The midpoints of the time intervals are

$$t_m = t_0 + (m + 1/2)\frac{L - r}{f_{\mathrm{s}}}$$

This scheme is frequently used in applications of spectral analysis, and is conveniently implemented in many software packages.

If desired for some reason, zero-padding could be put in after step 3, in which case a correction for the padding should be included in step 5. If the overlaps r and s are put equal, all data points except those at the very ends of the total time series will be used equally much in the analysis.

1.4 Wavelet Techniques

Wavelet analysis is a very rich field of techniques useful for many different applications, for example data analysis, theoretical electromagnetics, and data compression, as can be seen in any of the many texts on the subject [e.g. *Daubechies*, 1990; *Kaiser*, 1994; *Strang and Nguyen*, 1996]. We will here take a narrow-minded approach, only considering wavelet methods as an alternative to the Fourier methods for spectral estimations. Discussions of this aspect of wavelets can be found in the literature on applications to space data [e.g. *Holter*, 1995; *Lagoutte et al.*, 1992].

1.4.1 Morlet Wavelets

In Section 1.2.6, we found that the time-frequency plane may in principle be partitioned in many different ways. We now turn to the question of how to actually implement a partition of the type shown in panel (d) of Figure 1.2. The answer lies in wavelet analysis, which is unlike the traditional Fourier techniques in that it intrinsically relies on a time-frequency approach. The basic idea of wavelet analysis is to expand a signal in basis functions which are localised in time as well as frequency, so that they have the character of wave packets. This places wavelet methods somewhere between the Fourier techniques, where the basis functions $\exp(-2\pi \mathrm{i} f t)$ are sharp in frequency but completely spread out in time, and the pure time series representation, which offers perfect localisation in time but includes all frequencies[3].

[3] In practice, the basis functions of Fourier analysis are in fact localised in time by the data window, and the time series is localised in frequency by its finite sampling rate. However, all Fourier basis functions (all frequencies) are localised in the same time interval, which is not the case in wavelet analysis.

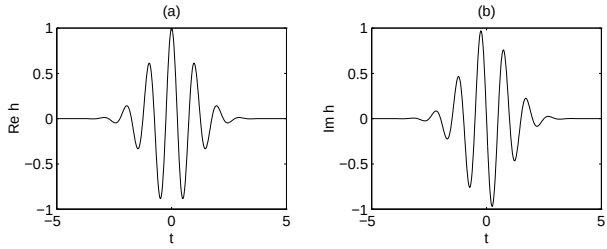

Figure 1.9: The Morlet mother wavelet, defined by equation 1.34, with $\omega_0 = 2\pi$. (a) Real part. (b) Imaginary part.

From a physical viewpoint, localisation in time as well as in frequency is an attractive perspective for the analysis of non-stationary signals. In particular, a basis consisting of localised packets of sinusoidal waves is appealing, as sinusoidal waves are the eigenmodes of a plasma. For the envelope of the wave packet, a Gaussian is a natural choice. As is well known from quantum mechanics, a Gaussian wave packet optimises localisation in both time and frequency as it is the only wave packet for which we get a '=' rather than a '≥' in the uncertainty principle $\delta f\, \delta t \geq 1$ [*Landau and Lifshitz*, 1977, p. 48]. This leads us naturally to the Morlet wavelet

$$h(t) = \exp(-t^2/2)\ \exp(-\mathrm{i}\,\omega_0\, t) \tag{1.34}$$

As ω_0 is a free parameter, equation 1.34 defines a family of functions. The value of ω_0 determines the number of oscillations in a wave packet, and has to be sufficiently large for 1.34 to be useful as a wavelet, due to problems with non-vanishing mean value of the real part. For $\omega_0 = 2\pi$, which we will use here for reasons to be seen later (equation 1.36), this error is negligible in practice. In space plasma applications, our choice $\omega_0 = 2\pi$ has sometimes been used [e.g. *Dudok de Wit et al.*, 1995], although $\omega_0 = 5$ seems more common [e.g. *Holter*, 1995; *Lagoutte et al.*, 1992], probably because it is close to the value $\pi\sqrt{2/\ln 2} \approx 5.34$ originally used by *Morlet et al.* [1982]. For Morlet's ω_0, the amplitude decreases to half its maximum in one period of the wave.

Many other wavelet families than the Morlets are possible, and several classes can be found in any text on wavelet methods. For our purpose, which is the study of spectral properties of non-stationary time series, we take the view that the Morlet wavelet, with its clear physical interpretation as a modulated sinusoidal oscillation and good properties of localisation in frequency as well as in time, is the natural choice.

By stretching and translation of a wavelet like 1.34, called a mother wavelet, we can get a whole set of wavelets of the same shape, known as daughter wavelets. It is customary to denote the stretching and translation by two parameters a and τ known as scale (or dilation) and translation, respectively. The daughter wavelets $h_{a\tau}(t)$ are then written as

$$h_{a\tau}(t) = \frac{1}{\sqrt{a}} h\left(\frac{t-\tau}{a}\right) \tag{1.35}$$

The concept of scale is used here instead of the concept of frequency, which most often is associated with sinusoidal functions. However, for the specific case of the Morlet wavelets,

which are based on sinusoidal functions, it is reasonable to replace the scale a by the frequency $f = 1/a$, so we write the daughter wavelets as

$$h_{f\tau}(t) = \sqrt{f}h(f\,(t-\tau)) = \sqrt{f}\,\exp\left(-\frac{f^2(t-\tau)^2}{2}\right)\,\exp\left(-2\,\pi\,\mathrm{i}\,f\,(t-\tau)\right) \quad (1.36)$$

This explains our choice $\omega_0 = 2\pi$: the Morlet wavelets become Gaussian envelopes of a carrier wave with frequency f.

The transformation above ensures that all daughter wavelets will look like their mother wavelet. Irrespective of f and τ, equally many periods of the oscillation will fit into the packet.

We can now define the Morlet wavelet transform (MWT) of the signal $u(t)$ by

$$C(\tau, f) = \int u(t)\, h^*_{f\tau}(t)\,\mathrm{d}t \quad (1.37)$$

In principle, the limits of integration should be $\pm\infty$. However, as the wavelets $h_{f\tau}(t)$ are localised in time, little error is introduced by integrating over a finite time interval. Finally, by going from integrals to sums in a fashion similar to how we introduced the discrete Fourier transform in Section 1.2.4, we make possible the practical evaluation of wavelet coefficients.

1.4.2 Wavelets and Fourier Methods

The Morlet wavelets 1.36 can be written on the form

$$h_{f\tau}(t) = A\, w_\tau(t, f)\,\exp(-2\pi\mathrm{i}ft) \quad (1.38)$$

where

$$w_\tau(t, f) = \exp\left(-\frac{f^2(t-\tau)^2}{2}\right) \quad (1.39)$$

and

$$A = \sqrt{f}\,\exp(2\pi\mathrm{i}f\tau) \quad (1.40)$$

Hence, the Morlet wavelet transform 1.37 may be written as

$$C(\tau, f) = A^* \int w_\tau(t, f)\, u(t)\,\exp(2\pi\mathrm{i}ft)\,\mathrm{d}t \quad (1.41)$$

When we discussed the use of window functions for Fourier methods in Section 1.3.3, we only considered sampled time series. However, for a continuous signal $u(t)$ to which we have applied a window function centred at $t = \tau$, denoted $W_\tau(t)$, the Fourier integral 1.3 is

$$\tilde{u}_\tau(f) = \frac{1}{T}\int W_\tau(t)\, u(t)\,\exp(2\pi\mathrm{i}ft)\,\mathrm{d}t \quad (1.42)$$

Now assume the window function is a Gaussian. Apart from the factors in front of the integrals, the only difference between equations 1.41 and 1.42 then is that the window function $w_\tau(t, f)$ in 1.41 depends on frequency as well as on time, while the window

$W_\tau(t)$ in 1.42 is the same for all frequencies. However, for any frequency f_0, we can always choose the Gaussian window so that $W_\tau(t) = w_\tau(t, f_0)$. At this frequency, it is possible to interpret the Morlet wavelet transform as a Fourier transform with a Gaussian window.

This result is useful for our understanding of wavelet methods. It enables us to apply many results and methods of Fourier analysis also to the wavelet transforms. We list a few of them here:

1. PSD estimation. The similarity of equations 1.41 and 1.42 indicates how to obtain a PSD estimate with physically meaningful normalisation with wavelet methods (equation 1.43).

2. Phase. We can also construct a wavelet phase spectrum corresponding to the Fourier phase spectrum 1.11.

3. Detrending. Frequency components below the lowest resolvable will affect the wavelet methods as well as the Fourier methods, so detrending (Section 1.3.2) may be useful.

4. Windowing is obviously inherent in the wavelet transform.

5. Averaging. For a random signal, the wavelet based PSD will also show statistical fluctuations, which in principle may be quenched by averaging in time, at the expense of temporal resolution[4] (Section 1.3.4).

6. Zero-padding is not a useful technique for wavelet analysis, since the zeroes cannot be padded after the windowing (Section 1.3.6) in a wavelet transform, where the window is implicit in the basis wavelet itself.

For the PSD, by comparing the Morlet wavelet transform 1.41 to the Gaussian window Fourier transform 1.42, using the definition 1.8 of the power spectral density and correcting for the window (envelope) by the factor W_{ss} defined by 1.27, we conclude that the PSD definition for the MWT which gives PSD values equal to what we find for the corresponding Gaussian windowed DFT estimate is

$$S_u^{\mathrm{MWT}}(t, f) = \frac{2}{\sqrt{\pi}} |C(t, f)|^2 \tag{1.43}$$

With this definition, the PSD values derived from Fourier analysis and wavelet transforms are comparable, in the sense that the PSD estimates for the shaded squares in panels (c) and (d) of Figure 1.2 are equal.

1.4.3 Continuous Wavelet Transform (CWT)

The wavelet transform is based on wave packets where the relation between frequency and packet width is constant. Hence, it naturally provides a means to implement the octave band analysis suggested by the division of the time-frequency plane in panel (d) of

[4] This somewhat unconventional viewpoint is further discussed below in Section 1.4.4.

Figure 1.2. As the t-f plane cannot be divided into more rectangles than half the number of samples in the time series, evaluation of one wavelet coefficient for each of these rectangles should give a complete description of the PSD.

In practical applications of wavelet transforms to spectral problems, one often evaluates more wavelet coefficients than is actually needed. This is known as doing a continuous wavelet transform (CWT), where the word continuous signifies that we evaluate 1.37 at freely chosen f and t. However, one should note that the time-frequency resolution of the CWT will be as depicted in panel (d) of Figure 1.2, even though the CWT may be evaluated at many points. We cannot get around the restrictions of the uncertainty relation 1.22. ⇒1.1

In the same way, one may also define a continuous Fourier transform, which simply amounts to extending the definition 1.14 to non-integer values of $n = f/f_s$, and letting the data windows overlap arbitrarily much. For an inverse transform back to the time domain, the non-integer values of n are of course entirely superfluous: no new information is gained by evaluating the DFT or the MWT for more time-frequency locations than $N/2$, and hence these extra coefficients never enter in an inverse transform. In fact, this sort of continuous Fourier transform is never used in practice, as it cannot be implemented with the FFT scheme. If one desires a smoother evaluation of the PSD, zero padding (Section 1.3.6) is used instead.

1.4.4 Comparing WFT and MWT: an Example

An example of a comparison of wavelet and Fourier methods for PSD calculation is shown in Figure 1.10. Panel (a) shows a time series of the electric field sampled at 32 000 samples/s by the wave instrument on the Freja satellite. The data shows lower hybrid waves of a bursty or modulated character [*Eriksson et al.*, 1994]. Panels (b)–(e) display different time-frequency spectral representations of the signal in (a). The total number of samples is $N = 13000$. The methods used for the spectral estimation all have different time resolution, and only the time interval for which all yield results is shown. For all the spectra, the number of displayed bins in the time-frequency plane are much higher than the number $N/2$ which defines the real information content in the spectra (Section 1.2.6), giving a smoother variation in time and/or frequency.

Panel (b) shows a Fourier time-frequency analysis based on averages of eight 64-point DFTs overlapping by 50%. The effective time resolution therefore is 512 samples or 16 ms, but the PSDs shown are separated only by one sample, giving a smoother time variation in the plot. As discussed in Section 1.3.4, the use of the averaging method is motivated if we interpret the signal as an almost stationary random process, whose slowly varying statistical parameters we want to estimate. The result of the averaging is to reduce the variations between consecutive spectra, and the spectrum is very clean. On the other hand, the resolution in the time-frequency plane goes down by a factor of eight as indicated by equation 1.29, and little trace of the bursty nature of the time series in panel (a) is retained.

In panel (c), we show a Fourier-based PSD estimate with the same time resolution (512 samples) as in (a). Here, no averaging has been used, and the effect of random fluctuations is therefore stronger. This results in larger variations within spectra, as is seen by the many horizontal structures in the plot. The frequency resolution in (c) is much better than in (b), and splittings of the spectral peak around 4 kHz can be clearly seen. On the other

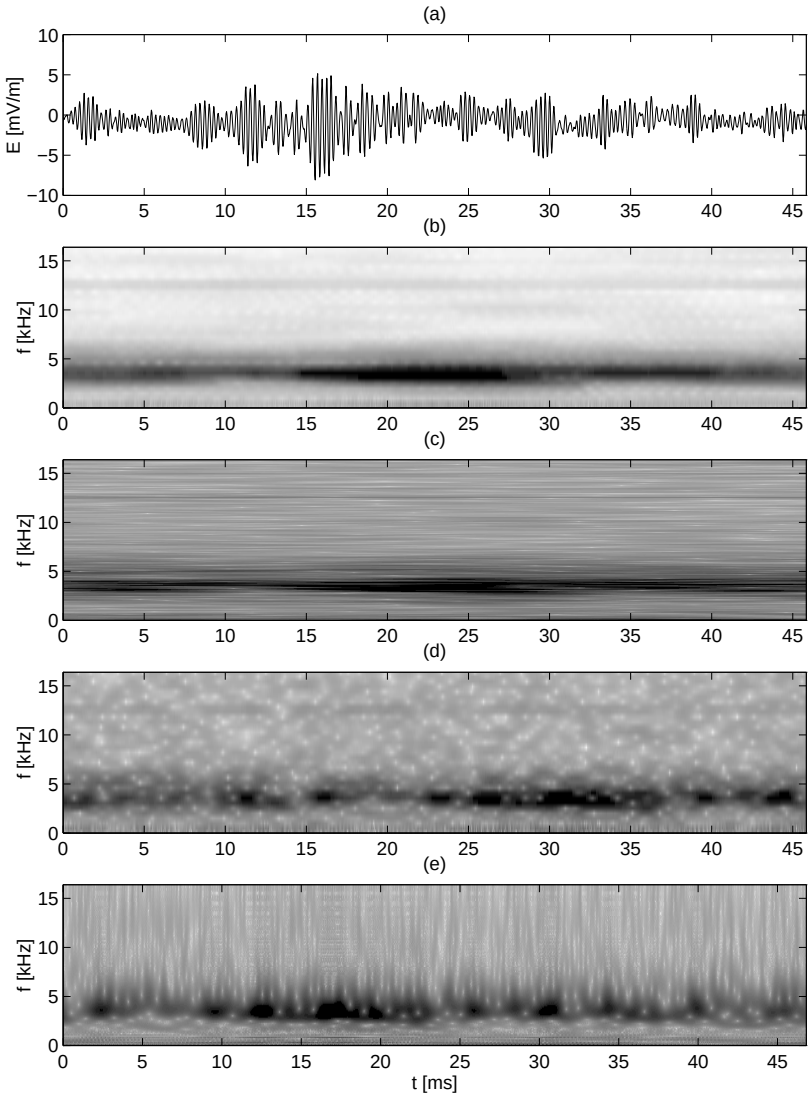

Figure 1.10: A time series (a) with time-frequency spectra obtained by different methods: (b) Averaged DFT analysis, 64 points, 8 averages. (c) High frequency-resolution DFT analysis, 512 points. (d) High time-resolution DFT analysis, 64 points. (e) Morlet wavelet analysis.

hand, little time variation can be discerned. The modulated nature of the time series is represented by frequency splittings rather than as time variations, as discussed in the text on page 13 and exemplified in Figure 1.3.

Panel (d) displays the result of a 64-point Fourier analysis without averaging. Hence, the frequency resolution is as for panel (b), while the time resolution is much better. Variations of the amplitude of the time series are now reflected in the time domain, as variations of the PSD with time. While the structures in (c) were almost horizontal, they are here fairly isotropic.

Finally, panel (e) shows the PSD based on Morlet wavelet analysis. The modulational pattern of the signal is represented as time variations of the PSD, as in panel (d). The structures to be seen in this plot are horizontal at the lowest frequencies (few kHz), and vertical at higher frequencies (above 5 kHz). This behaviour reflects that in the wavelet analysis, the time-frequency plane is partitioned as in panel (d) of Figure 1.2. This results in characteristic flame-like patterns in the plot: like the flames in a fire, the horizontal scale sizes are smaller at the top than in the base. These flame patterns is analogous to the large variance in PSD estimates by DFT methods, discussed in Section 1.3.4. They may very well reflect the random variations of a stationary process, in which case it may be justified to remove them by averaging in time. This is not normally discussed in wavelet texts, as one of the fundamental virtues of wavelet analysis is their usefulness for studying non-stationary phenomena. However, as such flames will turn up even if analysing a stationary process like Gaussian random noise, one should be careful not to put too much physical meaning in their appearance, and sometimes one may consider averaging in time.

All the Fourier-based plots (b, c, d) more or less clearly show a narrow spectral line near 12 kHz. This component cannot be seen in the wavelet plot (e). The reason for this is obvious from a glance at Figure 1.2. The partitioning of the time-frequency plane in wavelet analysis is such that all spectral details in the upper half plane (above $f_s/4$) are completely lost. In this particular case, the origin of the 12 kHz line is purely instrumental, so little harm is done by using the wavelet analysis here, but this is certainly not the general case.

Even though the plots in Figure 1.10 show continuous transforms evaluated at many more points than $N/2$, we see that the partitions of the t-f plane (c) and (d) in Figure 1.2 are inherent to the Fourier and wavelet methods, respectively, looking at the structuring of spectral detail. Evaluating the PSDs at very many points cannot hide these fundamental aspects of the methods, since there is no extra information gained by displaying more than $N/2$ points in the t-f plane. There is no way to fool relation 1.22.

1.4.5 Implementations of the Wavelet Transform

For some wavelet families, there are very efficient computation schemes. For others, including the Morlet wavelets, there are no algorithms as fast as the FFT routine of Fourier analysis. Discussions on algorithms can be found in standard texts, but fortunately these are already implemented in many commercial software packages, including the commonly used IDL and Matlab (as a separate toolbox), so you should not have to write them all yourself. Normalisation can be a problem when using such routines, since it is not always simple to deduce from the documentation what normalisation conventions are used, but the correspondence between Fourier and wavelet methods discussed in Section 1.4.2 can be used to derive formulas like 1.43 and for checking normalisation.

1.4.6 Wavelet Packet Analysis

As the wavelet analysis discussed above is based on wavelets with a similar number of wave periods in the wave packet for all carrier frequencies, it is naturally associated with the time-frequency partitioning shown in Figure 1.2d on page 12. However, it may very well be that the signal consists of a few transient pulses at long time scales, combined with stable narrow-band emissions at high frequencies. In that case, a partitioning corresponding to turning Figure 1.2d upside down could be advantageous. One could also think of a signal whose natural representation is in some completely unsymmetric division of the t-f plane, like what we get if we interchange the dashed rectangular areas in panels (c) and (d) of Figure 1.2. In terms of a Morlet wavelet framework, this would correspond to using basis wavelets 1.36 with varying values of ω_0, i.e. with a varying number of wave periods fitting into a wave packet.

Schemes for wavelet analysis along these lines are known as wavelet packet analysis [e.g. *Strang and Nguyen*, 1996]. This includes automatic search for a best basis, in which as much signal energy as possible is concentrated into as few t-f-boxes as possible. This has mainly found applications for purposes like image compression, but clearly holds an interesting potential for analysis of space plasma time series. A brief treatment of applications to geophysics, with references to theoretical work, is included in the review article by *Kumar and Foufoula-Georgiou* [1997].

1.5 Spectral Analysis of Multiple Signals

1.5.1 Cross-Spectral Analysis

Cross-spectral analysis of two time series is a means for looking at the relations between individual spectral components of two data records, usually sampled simultaneously. We define the cross-spectral density (CSD) of the two time series $u[j]$ and $v[j]$, with DFTs $\tilde{u}[n]$ and $\tilde{v}[n]$, respectively, by

$$G_{uv}[n] = \frac{2N}{f_s} \tilde{u}^*[n]\, \tilde{v}[n] \tag{1.44}$$

where * indicates complex conjugation. With this normalisation, the PSD becomes a special case of the cross-spectral density,

$$S_u[n] = G_{uu}[n] \tag{1.45}$$

In this chapter, we have confined ourselves to the case when u and v are signals of the same kind sampled at different locations, for example by different satellites. Cross-spectral methods are also very useful for the analysis of several different signals sampled at one location in space. One then forms a spectral matrix consisting of the cross-spectra of the signals, whose diagonal elements are the PSDs of the signals. This matrix can then be used for the study of wave properties like wave normal directions and field polarisation. A nice example is the classical study by *Means* [1972], and one may note that the STAFF instrument on the Cluster mission implements on-board spectral matrix analysis of five field components [*Cornilleau-Wehrlin et al.*, 1997]. The multi-spacecraft filtering technique presented in Chapter 3 is a further development, where signals from each of several measurement locations are included.

Unlike the PSD, the cross-spectral density 1.44 of real signals will in general be complex-valued. One introduces the coincident spectral density C_{uv} and the quadrature spectral density Q_{uv} as the real and imaginary parts of G_{uv}, respectively:

$$G_{uv}[n] = C_{uv}[n] + \mathrm{i}\, Q_{uv}[n] \tag{1.46}$$

The phase of the cross-spectral density is defined by

$$G_{uv}[n] = |G_{uv}[n]|\, \exp(\mathrm{i}\varphi_{uv}[n]) \tag{1.47}$$

so that

$$\tan \varphi_{uv}[n] = \frac{Q_{uv}[n]}{C_{uv}[n]} \tag{1.48}$$

By considering the definition 1.11 of the phase spectrum of a single signal, it follows that

$$G_{uv}[n] = |G_{uv}[n]| \exp(\mathrm{i}\varphi_{uv}[n]) = |\tilde{u}[n]|\, |\tilde{v}[n]|\, \exp(\mathrm{i}\{\varphi_v[n] - \varphi_u[n]\}) \tag{1.49}$$

Therefore, the phase of the cross-spectrum is the phase difference between the two signals.

To see the physics of the coincident and quadrature spectral densities defined by 1.46 above, we note that $C_{uv} = |G_{uv}| \cos \varphi_{uv}$ and $Q_{uv} = |G_{uv}| \sin \varphi_{uv}$. Thus C_{uv} and Q_{uv} represent the in-phase and out-of-phase contributions from the two signals to the total CSD. A simple example is the case where v is the voltage variation (in volts) in some system and u is the current in the same system (in ampères). In this case, C_{uv} and Q_{uv} are the spectra of effective and reactive power, respectively, in watts per hertz[5]. The impedance spectrum is $Z[n] = G_{uv}[n]/S_v[n]$ (in ohms), where the real part $R[n] = C_{uv}[n]/S_v[n]$ is the resistance spectrum, while the imaginary part $X[n] = Q_{uv}[n]/S_v[n]$ is the reactance spectrum, with contributions from the inductive and capacitive reactance spectra.

The implementation of cross-spectral methods by Fourier methods is straightforward. In addition, high-level data processing packages like IDL and Matlab include functions for phase spectra, coherence and other cross-spectral quantities. The application of cross-spectral analysis to space plasma data will be exemplified in Section 1.5.3. Before that, we must have a look at how to implement the definitions of the cross-spectral quantities and interpret the results.

1.5.2 Averaging and Coherence

From the definitions above, it is clear that as soon as we have the Fourier transforms of two signals, we can calculate a phase spectrum. However, just doing so and using the result to find for example a wavelength spectrum is not always a good idea. First, if we just consider the phase spectrum, there is obviously no information on signal strength in it, so frequencies with a clear signal are treated in the same way as frequencies only containing instrumental noise. Therefore, it is necessary to compare the phase spectrum to the CSD magnitude or to the PSDs of the signals, to see that there really is significant signal strength at the frequencies we are considering.

Second, how do we know that there really is any relation between the two signals at the frequency we are considering? The definitions above allow us the possibility of always

[5] Conventionally volt-ampères per hertz for the reactive power.

calculating the phase spectrum, irrespective of any causal relation between the signals. In the case of a stationary process, one can define a coherence spectrum,

$$\gamma_{uv}^2[n] = \frac{|G_{uv}[n]|^2}{S_u[n]\, S_v[n]} \tag{1.50}$$

which is useful in this context. At first this may seem strange: from the definition of G_{uv} 1.44 and the relation $S_u = G_{uu}$ 1.45, one may conclude that $\gamma_{uv}^2[n] \equiv 1$, regardless of the signals u and v. This is certainly true, unless the spectral estimates G_{uv}, S_u, and S_v are calculated using the averaging method of Section 1.3.4. If averaging is included in the calculation of the spectral densities, the value of γ_{uv}^2 can end up anywhere between zero and one. As the denominator in 1.50 is positive, the coherence is essentially determined by the constancy of the phase of the numerator, i.e. by the stability of the phase spectrum $\varphi_{uv}[n]$. If, for a certain frequency f_n, the phase spectrum changes rapidly between the segments included in the averaging process, the averaged G_{uv} for this frequency can take any complex value. In the case of random phase fluctuations and sufficiently many averages, the value of the coherence will come close to zero. The coherence function 1.50 therefore indicates the stability of the phase spectrum estimate. A low value of the coherence at some frequency implies that the phenomenon behind the phase spectrum estimate at this frequency cannot be considered correlated between the two measurement points and stationary in time. Hence, a usual prescription for cross-phase methods is to include averaging in the calculation and only use points which turn out to show values of γ_{uv}^2 higher than some threshold value.

It is important to note that the coherence concept is meaningful only for stationary signals. Still, the phase concept may be useful also for non-stationary signals, for example a wave packet seen only briefly in the spacecraft frame. If we have some reason to believe that there is a relation between two signals u and v, it is perfectly legitimate to examine their cross-spectrum even if the observation time is so short that no averaging and hence no coherence estimate can be done. This is the situation we encountered on page 20, where we discussed averaging. If we interpret our signals as produced by a stationary random process, averaging is useful, and the coherence spectrum is appropriate for determining the quality of the phase spectrum. On the other hand, if we interpret the signals as a unique record of some non-stationary phenomenon, averaging and hence the coherence function are irrelevant concepts. In this case, one must resort to other means for judging the reliability of the phase spectrum. Checking the signal intensity from the PSD or CSD is an obvious approach. It may also help to inspect the phase spectrum itself: if it is fairly smooth, its interpretation is safer than if it is wildly and randomly fluctuating with changing frequency.

For non-stationary signals, it may also be a good idea to use wavelet rather than Fourier methods as a basis for the cross-spectral analysis. We return to this question in Section 1.5.4.

1.5.3 Cross-Spectral Analysis: an Example

As a simple example of the use of cross-spectral analysis in space applications, we show an analysis of Langmuir probe current data from the Freja wave instrument in Figure 1.11. Panel (a) shows the relative variations of the current $\delta I/I$ collected by two

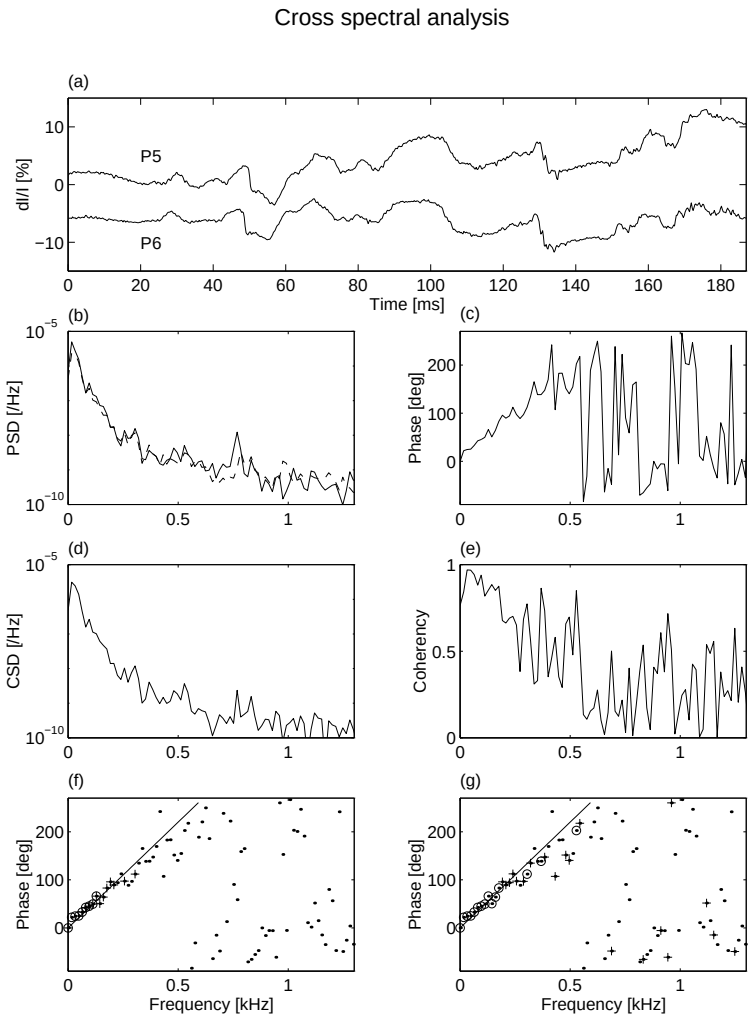

Figure 1.11: Langmuir probe data from the Freja wave instrument illustrating the use of cross-spectral analysis. All spectra calculated using averaging of five 256-point DFTs with 128 point overlap. (a) Time series of probe current fluctuations from two probes, sampled at 4 ksamples/s. (b) Power spectra of the P5 (solid) and P6 (dashed) signals. (c) Relative phase spectrum. (d) Magnitude of cross-spectral density. (e) Coherence spectrum. (f) Relative phase spectrum coded by CSD magnitude. Circles: $|\mathrm{CSD}| > 10^{-7}\,\mathrm{Hz}^{-1}$. Crosses: $10^{-8}\,\mathrm{Hz}^{-1} < |\mathrm{CSD}| < 10^{-7}\,\mathrm{Hz}^{-1}$. Dots: $|\mathrm{CSD}| < 10^{-7}\,\mathrm{Hz}^{-1}$. (g) Relative phase spectrum coded by coherence. Circles: $\gamma^2 > 0.75$. Crosses: $0.5 < \gamma^2 < 0.75$. Dots: $\gamma^2 < 0.5$.

positively biased spherical probes P5 and P6 mounted on 5.6 meter booms at opposite sides of the spacecraft. The curves have been shifted along the ordinate in order to separate them. These two time series thus constitute a multipoint measurement in space. The two curves are obviously very similar.

Figure 1.11(b) shows the power spectra of the two $\delta I/I$ signals. As expected, they are very similar, both showing a decay with increasing frequency. The phase and amplitude of the CSD are shown in panels (c) and (d), respectively, and the coherence as defined by 1.50 is plotted in panel (e). As the spectra in (b) are similar to each other, it is not surprising that (d) looks about the same as well. The coherence spectrum in (e) shows a decay with frequency reminiscent of the power spectra, although there are some coherence peaks at higher frequencies. Panels (f) and (g) show the phase spectrum again, this time with different symbols used for plotting values corresponding to different ranges of CSD and γ^2 (see caption).

For multipoint measurements, an important feature of the phase difference is that it provides information on wavelengths. Consider a wave field varying as $\exp(\mathrm{i}[k\,x - \omega\,t])$. Measurements at two points separated by a distance d will have a phase difference

$$\varphi = k\,d \tag{1.51}$$

which can be revealed by a cross-spectral analysis. Knowing the separation distance, the wave number k and hence the wavelength of the wave would be known. In the case of Figure 1.11, the points with high CSD and coherence at low frequencies in panels (f) and (g) lines up around a line $\varphi = a\,f = a'\,\omega$, where $a \approx 440°/\mathrm{kHz} \approx 7.7 \cdot 10^{-3}$ rad/Hz which gives $a' \approx 2.4 \cdot 10^{-3}$ s. Comparing this to the predicted phase difference 1.51 using the known probe separation $d = 11.2$ m, we find that this corresponds to a phase velocity

$$v_\phi = \frac{\omega}{k} = \frac{d}{a'} \approx 4.6\,\mathrm{km/s} \tag{1.52}$$

along the line of probe separation in the reference frame of the spacecraft. In this case, the satellite speed was 6.1 km/s and the angle between the probe separation distance and the spacecraft velocity was 40°, so the observed behaviour simply reflects the satellite flying through essentially stationary structures in the plasma. However, the principle for the calculations would be the same also for travelling waves, and the method have been used for several successful measurements of wavelengths in space. For further examples and discussions we refer to the review by *LaBelle and Kintner* [1989], to the application papers by for example *Holmgren and Kintner* [1990] or *Vago et al.* [1992], and to the interesting extension of this method to the estimation of frequency-wavenumber spectra by *Beall et al.* [1982].

1.5.4 Cross-Wavelet Analysis

Cross-spectral analysis does not have to be based on Fourier methods. The Morlet wavelet analysis discussed in Section 1.4 is also possible to use for phase spectrum estimations, which can be advantageous for non-stationary processes. This approach has quite recently been applied to frequency-wavenumber spectral estimation in space plasma data by *Dudok de Wit et al.* [1995] and *Pinçon et al.* [1997].

The advantage of wavelet methods is in fact more pronounced for the study of phase spectra than for the PSD estimations we discussed in Section 1.4. The reason for this is

that while the PSD is a positive definite additive quantity for which averaging is useful, the phase spectrum can have all signs, and the average of a changing phase has no physical meaning. The phase estimates we get from a Fourier implementation of cross-spectral analysis is based upon some fixed time interval, equal for all frequencies. Hence, these estimates are based on a few wave periods for a low frequency wave, but many wave periods for an oscillation at higher frequencies. We thus intrinsically put higher demands on the phase stability of waves at higher frequencies as compared to low-frequency fluctuations. For PSD estimates, this is no fundamental problem. Estimating the power of a changing signal simply gives the mean power in the time interval in question.

To illustrate this, we consider the idealised example in Figure 1.12. One artificial data record (dashed) consists of unit amplitude sinusoids at 0.1 Hz with some added noise, sampled at 1 sample/s. The other record (solid) is similar, but with a 180° phase shift halfway through the interval. The total interval of 192 points is divided into five overlapping 64 point sections, which are Hamming windowed before DFT analysis. The averaged PSDs over these five sections are shown in panel (b). The fact that one of the signals is not stationary and changes its phase during the interval does not change its PSD (solid curve) very much, although there is some broadening. Panels (c) and (d) show the phase and the amplitude of the CSD, and panel (e) the coherence spectrum. It is seen that the coherence at the signal frequency 0.1 Hz is very low, as it must be due to the changing phase of one of the signals. Finally, panels (f) and (g) show the phase spectra based only on the first and last 64 points of data, respectively. These tells us the true story: the phase at 0.1 Hz is 0 at the beginning of the record and 180° at the end.

If the signal to be analysed by cross-spectral methods is non-stationary but still narrow in its frequency content, a Fourier approach may work well. One then divides the time series into records of a length corresponding to a few periods of the dominant component, and can then keep track of the changing phase. However, if the non-stationary signal contains a wide range of frequencies, a Fourier based cross-spectral method is unlikely to provide useful phase information over all the frequency spectrum. A wavelet implementation, using wavelet coefficients 1.37 instead of the DFT in the definition of the cross-spectral density 1.44 is then likely a better choice. This unavoidably means losing spectral detail at high frequencies, as the wavelet decomposition of course gives less frequency resolution at high frequencies than does its Fourier counterpart.

1.6 Parametric Methods

The fundamental idea of these methods, which have names like AR (autoregressive), MA (moving average), and ARMA, are to estimate the power spectrum of signal by estimating the parameters of a model for the time series of the signal. Parametric methods are akin to the Fourier techniques in the sense that their division of the time-frequency plane is similar, but otherwise quite different in spirit. These methods may very well dominate in modern textbooks on spectral analysis [e.g. *Brockwell and Davis*, 1987; *Kay*, 1988; *Marple*, 1987], and sometimes one will hear that these methods gives "better" spectral estimates than classical Fourier methods. It is important to realise that this is true only if the process we are studying really is of the kind assumed in the parametric model. To illustrate this point we will briefly discuss one of the best known and most used approaches, the AR method of spectral estimation, which for a Gaussian random process is equivalent to the

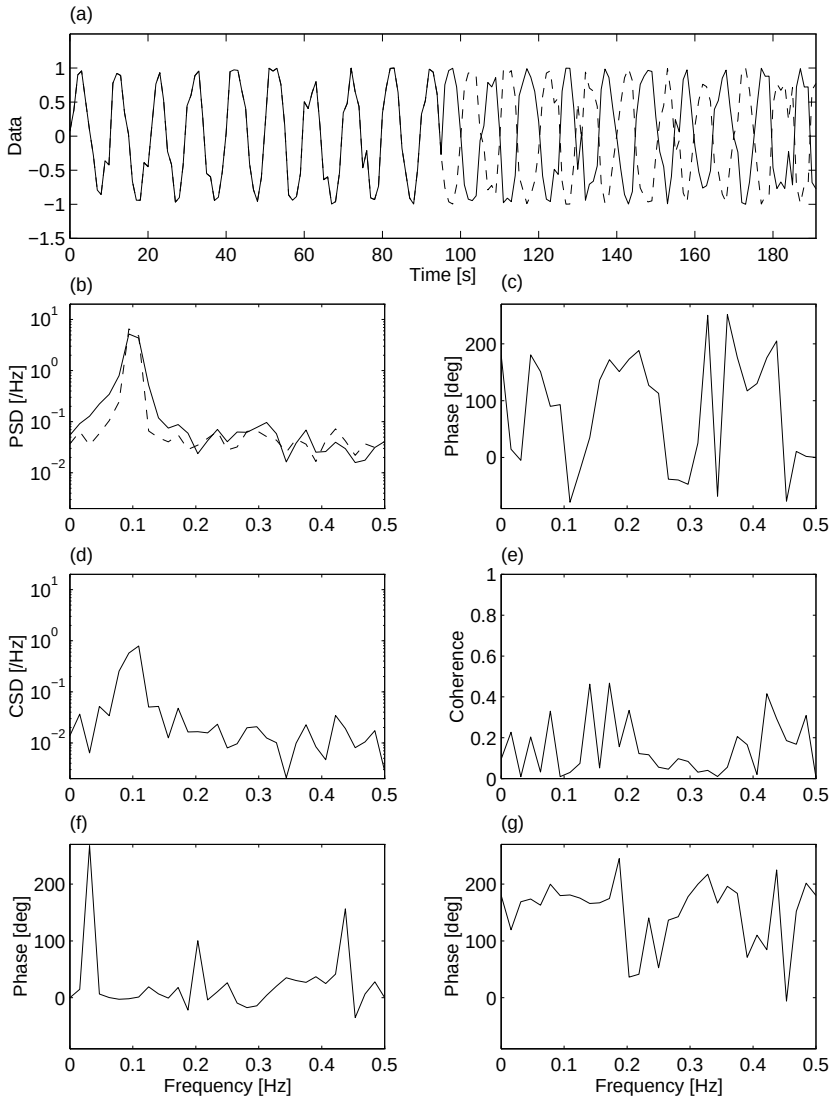

Figure 1.12: Synthetic signals with non-stationary relative phase, and cross-spectral analysis of these. autoregressive spectral methods. (a) Time series. (b) PSDs of the time series. (c) Phase spectrum based on the full record. (d) CSD of the full time series. (e) Coherence spectrum. (f) and (g) Phase spectra based on first and last thirds of the data record, respectively.

maximum entropy method (MEM) [*Ables*, 1974].

An AR model of order p models the time series as autoregressive, which means that the sample $u[j]$ is a linear function of the preceding samples:

$$u[j] = \sum_{k=0}^{p} a[k]\, u[j-k] + \eta[j] \tag{1.53}$$

where $\eta[j]$ represents uncorrelated noise. An AR model therefore is appropriate for a linear dynamical system. As linear systems are completely described by a set of eigenmodes with well defined frequencies, AR spectral estimation is good at representing spectra where the energy is concentrated in sharp peaks. Hence, if we have a situation where we have reason to expect well-defined eigenfrequencies, an AR algorithm is a good choice. For the type of problems facing the physicist studying processes in a turbulent space plasma, this is a rare case. Consider the example in Figure 1.13. A model PSD with a broad maximum superposed on a $1/f$ spectrum,

$$S(f) = \frac{1}{f} + 30\ \exp\left(-\left(\frac{f-0.2}{0.05}\right)^2\right) \tag{1.54}$$

was used for creating the time series in panel (a) as described in Section 1.2.3 with random phase spectrum $\phi[n]$. In panel (b), two sinusoids with amplitude 5 and frequencies 0.10 and 0.31 have been added. Panels (c) and (d) shows the model spectrum $S(f)$ (dashed) together with the spectral estimates by the Welch method (averaged and windowed DFT) and an AR method with $p = 10$. The Fourier estimate is the ragged curve which lies almost on the dashed model curve, while the smoother AR estimate mostly lies below. The advantage of the AR method is that it represents the two sinusoids in Panel (d) as narrow lines. However, if we are interested in anything else than these discrete frequencies, the Fourier method estimate clearly is better. The AR estimator tries to represent all spectral features by narrow peaks, including the broad peak around frequency 0.2 and the $1/f$ spectrum at low frequencies. When studying plasma processes in space, we are generally interested in the full spectrum, and the AR method is therefore usually not optimal. The same can be said for the application of other parametric spectral methods, although there of course are exceptions. An example where parametric methods indeed are appropriate is the search for signatures of solar eigenoscillations in the solar wind plasma by *Thomson et al.* [1995]. The solar oscillations are small-amplitude eigenmodes of a system not much perturbed by its environment, like a ringing bell, and hence it is reasonable to expect that the signatures of these oscillations can be described as an AR process. The eigenfrequencies carry information on the sun, and estimating them as exactly as possible is a task not suited for the Fourier and wavelet methods we emphasise in this chapter.

1.7 Final Comments

When doing spectral analysis of time series data from several spacecraft, two fundamental approaches suggest themselves. The first is to apply spectral analysis to the time series from each spacecraft separately, and then compare the spectral data. This is appropriate for the study of how processes on a much shorter length scale are modulated

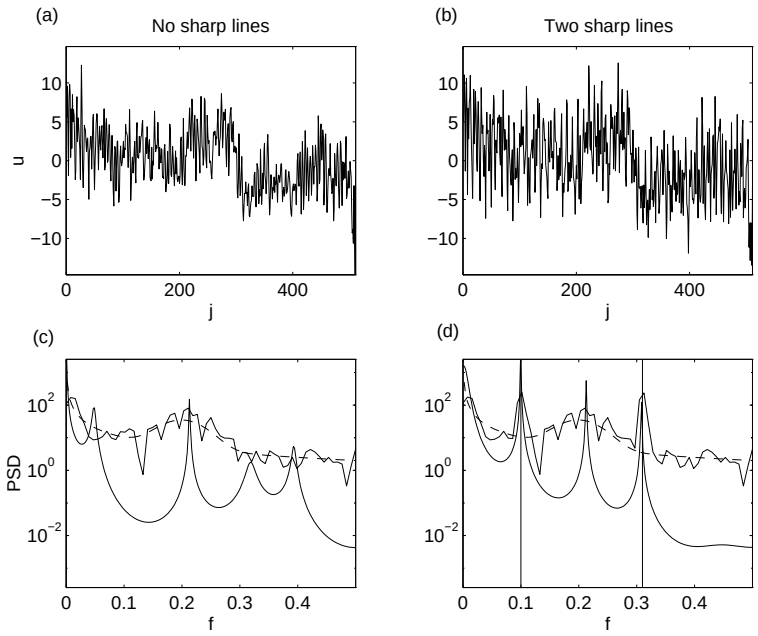

Figure 1.13: Synthetic signal and spectra calculated with Fourier and autoregressive spectral methods. (a) Signal constructed from the spectrum 1.54. (b) Two sinusoids at frequencies 0.10 and 0.31 added. (c) True spectrum of the signal in (a) (dashed), AR estimate (solid, sharply peaked) and DFT estimate (solid, following the dashed line). (d) As (c) but for the signal in (b).

at larger scales. The alternative approach, of an intrinsical multipoint nature, is to do spectral analysis of time series from two or more spacecraft together. The cross-spectral techniques described above in Section 1.5 belong to this class of methods, suitable for studying physical phenomena whose shortest spatial scale is on the order of the spacecraft separation distance. For another prime example of such a method, we refer the reader to the ***k***-filtering technique described in Chapter 3 of this book.

Bibliography

Ables, J. G., Maximum entropy spectral analysis, *Astron. Astrophys. Suppl. Series.*, **15**, 383–393, 1974.

Beall, J. M., Kim, Y. C., and Powers, E. J., Estimation of wavenumber and frequency spectra using fixed probe pairs, *J. Appl. Phys.*, **53**, 3933–3940, 1982.

Bendat, J. S. and Piersol, A. G., *Random data: Analysis and measurement procedures*, John Wiley & sons, 1971.

Brockwell, P. J. and Davis, R. A., *Time series: theory and methods*, Springer, 1987.

Brook, D. and Wynne, R. J., *Signal processing*, Edward Arnold, London, 1988.

Champeney, D. C., *Fourier transforms and their physical applications*, Academic Press, 1973.

Cornilleau-Wehrlin et al., The Cluster spatio-temporal analysis of field fluctuations (STAFF) experiment, *Space Sci. Rev.*, **79**, 107–136, 1997.

Daubechies, I., The wavelet transform, time-frequency localization and signal analysis, *IEEE Trans. Inform. Theory*, **36**, 961–1005, 1990.

Dudok de Wit, T., Krasnosel'skikh, V. V., Bale, S. D., Dunlop, M. W., Lühr, H., Schwartz, S. J., and Wolliscroft, L. J. C., Determination of dispersion relations in quasi-stationary plasma turbulence using dual satellite data, *Geophys. Res. Lett.*, **22**, 2653–2656, 1995.

Eriksson, A. I., Holback, B., Dovner, P. O., Boström, R., Holmgren, G., André, M., Eliasson, L., and Kintner, P. M., Freja observations of correlated small-scale density depletions and enhanced lower hybrid waves, *Geophys. Res. Lett.*, **21**, 1843–1846, 1994.

Glassmeier, K.-H. and Motschmann, U., Comments on time series analysis, in *Proceedings of Cluster Workshops, Braunschweig, 28–30 Sep. 1994, Toulouse, 16–17 Nov. 1994*, ESA SP–371, pp. 7–14, European Space Agency, Paris, France, 1995.

Harris, F. J., On the use of windows for harmonic analysis with the discrete Fourier transorm, *Proc. IEEE*, **66**, 51–83, 1978.

Holmgren, G. and Kintner, P. M., Experimental evidence of widespread regions of small-scale plasma irregularities in the magnetosphere, *J. Geophys. Res.*, **95**, 6015, 1990.

Holter, Ø., Wavelet analysis of time series, in *Proceedings of Cluster Workshops, Braunschweig, 28–30 Sep. 1994, Toulouse, 16–17 Nov. 1994*, ESA SP–371, pp. 43–50, European Space Agency, Paris, France, 1995.

Jenkins, G. M. and Watts, D. G., *Spectral analysis and its applications*, Holden Day, 1968.

Kaiser, G., *A friendly guide to wavelets*, Birkhäuser, 1994.

Kay, S. M., *Modern spectral estimation*, Prentice Hall, Englewood Cliffs, 1988.

Kumar, P. and Foufoula-Georgiou, E., Wavelet analysis for geophysical applications, *Rev. Geophys.*, **35**, 385–412, 1997.

LaBelle, J. and Kintner, P. M., The measurement of wavelength in space plasmas, *Rev. Geophys.*, **27**, 495, 1989.

Lagoutte, D., Cersier, J. C., Plagnaud, J. L., Villain, J. P., and Forget, B., High-latitude ionospheric electrostatic turbulence studied by means of the wavelet transform, *J. Atmos. Terr. Phys.*, **54**, 1283–1293, 1992.

Landau, L. D. and Lifshitz, E. M., *Quantum Mechanics (Non-Relativistic Theory), Course of Theoretical Physics vol. 3, 3rd edition*, Pergamon, 1977.

Marple, S. L., *Digital spectral analysis*, Prentice-Hall, Englewood Cliffs, 1987.

Means, J. D., The use of the three-dimensional covariance matrix in analyzing the polarization properties of plane waves, *J. Geophys. Res.*, **77**, 5551–5559, 1972.

Morlet, J., Arens, G., Fourgeau, I., and Giard, D., Wave propagation and sampling theory, *Geophysics*, **47**, 203–236, 1982.

Pinçon, J. L., Kintner, P. M., Schuck, P., and Seyler, C. E., Observation and analysis of lower hybrid solitary structures as rotating eigenmodes, *J. Geophys. Res.*, **102**, 17 283–17 296, 1997.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., *Numerical recipies in C*, Cambridge University Press, 2nd ed., 1992.

Rönnmark, K., Quantitative methods for waves in space plasmas, *Space Sci. Rev.*, **54**, 1–73, 1990.

Strang, G. and Nguyen, T., *Wavelets and filter banks*, Wellesley-Cambridge Press, Welles-

ley, MA., 1996.

Theiler, J., Eubank, S., Longtin, A., Galdrikan, B., and Farmer, J. D., Testing for nonlinearity in time series: the method of surrogate data, *Physica D*, **58**, 77–94, 1992.

Thomson, D. J., Maclennan, C. G., and Lanzerotti, L. J., Propagation of solar oscillations through the interplanetary medium, *Nature*, **376**, 139–144, 1995.

Vago, J. L., Kintner, P. M., Chesney, S. W., Arnoldy, R. L., Lynch, K. A., Moore, T. E., and Pollock, C. J., Transverse ion acceleration by localized lower hybrid waves in the topside auroral ionosphere, *J. Geophys. Res.*, **97**, 16 935–16 957, 1992.

Welch, P. D., The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Trans. Audio. Electroacoust.*, **AU-15**, 70–73, 1967.